DUKE MATHEMATICAL JOURNAL

EDITED BY

VOLUME 8

1941

DUKE UNIVERSITY PRESS
DURHAM, N. C.

0

CONTENTS

Volume 8, 1941

THE GROWTH OF THE SOLUTIONS OF A DIFFERENTIAL EQUATION

By N. Levinson

1. It has been shown[1] that the solutions of the differential equation

$$\frac{d^2x}{dt^2} + \phi(t)x = 0 \tag{1.1}$$

are bounded as $t \to \infty$ if there exists a constant $a > 0$ such that

$$\int_0^\infty | \phi(t) - a | \, dt < \infty. \tag{1.2}$$

It has also been shown[2] that if

$$\phi(t) = a + O\left(\frac{1}{t}\right), \tag{1.3}$$

the solutions need not be bounded. Here we shall go further in this direction and show that (1.2) is actually a best possible condition. We shall also show that if (1.2) is satisfied the solutions of (1.1) are not only bounded, but also resemble the solutions of the differential equation $x'' + ax = 0$ in another sense. Namely, if (1.2) is satisfied, then any solution of (1.1) satisfies also

$$\limsup_{t \to \infty} | x(t) | > 0.$$

In fact what we shall show is that the rapidity with which the solutions of (1.1) can grow and the rapidity with which they can tend to zero both depend on the growth of $\alpha(t)$, where

$$\alpha(t) = \int_0^t | \phi(t) - a | \, dt. \tag{1.4}$$

Thus the results for (1.2) will be a particular case of (1.4) where $\alpha(t)$ is bounded, and in this sense the first result we shall prove, Theorem I, is a generalization of the result of Fukuhara and Nagumo.

Received September 16, 1940.

[1] Fukuhara and Nagumo, *On a condition of stability for a differential equation*, Proc. Imp. Acad. of Japan, vol. 6(1930), pp. 131–132.

[2] O. Perron, *Über ein vermeintliches Stabilitätskriterium*, Göttinger Nachrichten, Math. Phys. Klasse, 1930, pp. 28–29, equation (6).

THEOREM I.[3] *If $x(t)$ satisfies the real differential equation* (1.1), *then*

$$x(t) = O(\exp [\tfrac{1}{2}a^{-\frac{1}{2}}\alpha(t)]), \tag{1.5}$$

where $\alpha(t)$ is defined as in (1.4).

Clearly if $\alpha(t)$ is bounded, it follows from (1.5) that $x(t)$ is bounded.

The question now arises as to how good an appraisal (1.5) is. We shall show that it fails in being a best possible result by at most a constant in the exponent of e. This is contained in the following result.

THEOREM II. *Let $\alpha(t)$ be a monotone increasing function such that $\alpha'(t) = o(1)$ as $t \to \infty$. Then there exists a $\phi(t)$ such that for large t*

$$\int_0^t | \phi(t) - a | \, dt \leqq \alpha(t) \tag{1.6}$$

and such that $x'' + \phi(t)x = 0$ has a solution $x(t)$ which satisfies

$$\limsup_{t\to\infty} \frac{\log | x(t) |}{\alpha(t)} \geqq \frac{1}{\pi}. \tag{1.7}$$

In particular it follows from (1.7) that condition (1.2) is a best possible condition for the boundedness of solutions of (1.1).

Next we shall show that the growth of $\alpha(t)$ determines not only how large a solution of the differential can become, but how small it can become.

THEOREM III. *If $x(t)$ satisfies* (1.1), *and if $\alpha(t) = O(t)$ as $t \to \infty$, then*

$$\limsup_{t\to\infty} | x(t) | \exp [\tfrac{1}{2}a^{-\frac{1}{2}}\alpha(t)] > 0, \tag{1.8}$$

where $\alpha(t)$ is defined as in (1.4).

Finally as with Theorem I, it can be shown that (1.8) fails in being a best possible result by at most a constant in the exponent of e as indicated in the following result.

THEOREM IV. *In Theorem* II (1.7) *can be replaced by*

$$\limsup_{t\to\infty} \frac{\log | x(t) |}{\alpha(t)} \leqq -\frac{1}{\pi}. \tag{1.9}$$

It will turn out that $\phi(t)$ as given in the proof of Theorems II and IV will be only piecewise continuous. In order to show that it is (1.4) that is relevant and not the local behavior of $\phi(t)$, we shall in §4 prove Theorems II and IV for a continuous $\phi(t)$, in fact for a $\phi(t)$ differentiable any finite number of times.

[3] In this and the theorems that follow, we shall assume that $\phi(t)$ is piecewise continuous and that $x(t)$ is a solution of (1.1) over $(0, \infty)$ if it satisfies (1.1) at all points of continuity of $\phi(t)$ and if $x(t)$ and $x'(t)$ are continuous over $(0, \infty)$.

2. Here we shall prove Theorems I and III:

Proof of Theorem I. Here $x(t)$ satisfies

$$x'' + \phi(t)x = 0, \tag{2.0}$$

and $\alpha(t)$ is defined as in (1.4). With no restriction we can assume that $x(t)$ is real. Clearly (2.0) can be written as

$$x'' + ax = (a - \phi(t))x. \tag{2.1}$$

Multiplying (2.1) by x' and defining

$$E(t) = \tfrac{1}{2}(x^2 + x'^2),$$

(2.1) becomes

$$\frac{dE}{dt} = (a - \phi(t))xx'. \tag{2.2}$$

But

$$| xx' | \leqq \tfrac{1}{2}(a^{-\frac{1}{2}} x^2 + a^{\frac{1}{2}} x'^2) = a^{-\frac{1}{2}} E(t).$$

Thus (2.2) can be written as

$$\frac{dE}{dt} \leqq a^{-\frac{1}{2}} \,|\, a - \phi(t) \,|\, E(t).$$

Or

$$\frac{dE}{E} \leqq a^{-\frac{1}{2}} \,|\, a - \phi(t) \,|\, dt. \tag{2.3}$$

Thus

$$\int_0^t \frac{dE}{E} \leqq a^{-\frac{1}{2}}\alpha(t).$$

Or

$$E(t) = O(\exp [a^{-\frac{1}{2}}\alpha(t)]).$$

From the definition of $E(t)$ it follows at once that

$$x(t) = O(\exp [\tfrac{1}{2}a^{-\frac{1}{2}}\alpha(t)]).$$

This completes the proof.

Proof of Theorem III. In much the same way as (2.2) yields (2.3), it also yields

$$\frac{dE}{E} \geqq -\, a^{-\frac{1}{2}} \,|\, a - \phi(t) \,|\, dt. \tag{2.4}$$

If we integrate, it follows that

$$\log E \geqq -\, a^{-\frac{1}{2}}\alpha(t) + \log C$$

for some constant C. This gives

$$E(t) \geqq C \exp[-a^{-\frac{1}{2}}\alpha(t)]$$

or

$$\left(\frac{dx}{dt}\right)^2 + ax^2 \geqq 2C \exp[-a^{-\frac{1}{2}}\alpha(t)]. \tag{2.5}$$

Over $n \leqq t \leqq n + 1$ let the maximum value of $|x(t)|$ be denoted by x_n and let a point where this value is taken on be t_n. Let a point where the minimum of $|x'(t)|$ over $n \leqq t \leqq n + 1$ is taken be t_n'. Then either $x'(t_n') = 0$ or else $x'(t)$ does not change its sign over $(n, n + 1)$. Suppose the latter is the case. Clearly

$$x(n + 1) = x(n) + \int_n^{n+1} x'(t)\, dt.$$

Thus

$$|x(n + 1)| + |x(n)| \geqq \left|\int_n^{n+1} x'(t)\, dt\right|.$$

Since $x'(t)$ does not change its sign in $(n, n + 1)$, this inequality yields

$$|x(n + 1)| + |x(n)| \geqq |x'(t_n')|.$$

In turn this yields

$$2x_n \geqq |x'(t_n')|. \tag{2.6}$$

This last inequality holds if $x'(t_n') = 0$. Thus it is always true.

Clearly

$$x'(t_n) = x'(t_n') + \int_{t_n'}^{t_n} x''(t)\, dt.$$

Thus

$$|x'(t_n)| \leqq |x'(t_n')| + \int_n^{n+1} |x''(t)|\, dt.$$

Or using the differential equation (1.1), we get

$$|x'(t_n)| \leqq |x'(t_n')| + \int_n^{n+1} |\phi(t)x(t)|\, dt.$$

If we use (2.6), it follows from the above inequality that

$$|x'(t_n)| \leqq 2x_n + x_n \int_n^{n+1} |\phi(t)|\, dt.$$

Using this in (2.5) at $t = t_n$ gives

$$x_n^2\left[2 + \int_n^{n+1} |\phi(t)|\, dt\right]^2 + ax_n^2 \geqq 2C \exp[-a^{-\frac{1}{2}}\alpha(t_n)]. \tag{2.7}$$

Since $\alpha(t) = O(t)$, it follows that, for an infinite number of n, $\int_n^{n+1} | \phi(t) | \, dt$ is bounded. Thus for an infinite number of n, (2.7) gives

$$x_n^2 \geqq C_1 \exp [-a^{-\frac{1}{2}}\alpha(t_n)],$$

where C_1 is some positive constant. The above inequality can be written as

$$\limsup_{t\to\infty} | x(t) | \exp [\tfrac{1}{2}a^{-\frac{1}{2}}\alpha(t)] > 0.$$

This completes the proof of Theorem III.

3. Here we shall give the proofs of Theorems II and IV. It will be convenient to assume that in (1.6) $a = 1$. This can always be done by changing the variable t to $a^{-\frac{1}{2}}t$ in the differential equation. Thus in place of (1.6) we have

$$\int_0^t | \phi(t) - 1 | \, dt \leqq \alpha(t). \tag{3.1}$$

Also with no real restriction, since $\alpha'(t) = o(1)$ and since our conclusions concern $t \to \infty$, we can assume that $\alpha(0) = 0$ and that

$$\alpha'(t) < \tfrac{1}{10}, \qquad 0 < t < \infty.$$

Let $\epsilon > 0$ be a small number and let

$$\beta(t) = \frac{1-\epsilon}{\pi}\alpha(t).$$

Let

$$\begin{aligned} &t_0 = 0, \qquad t_2 = \tfrac{1}{2}\pi\{1 + \exp [\beta(\tfrac{1}{2}\pi)]\}, \\ &t_{2n+1} - t_{2n} = \tfrac{1}{2}\pi \qquad (n = 0, 1, 2, \cdots), \\ &t_{2n} - t_{2n-1} = \tfrac{1}{2}\pi \exp [\beta(t_{2n-1}) - \beta(t_{2n-3})] \quad (n = 2, 3, \cdots). \end{aligned} \tag{3.2}$$

This defines t_n for all $n \geqq 0$.

LEMMA 1. *As* $n \to \infty$,

$$t_{n+1} - t_n = \tfrac{1}{2}\pi + o(1).$$

Proof. Adding the lower two equations of (3.2) gives

$$t_{2n+1} - t_{2n-1} = \tfrac{1}{2}\pi \{1 + \exp [\beta(t_{2n-1}) - \beta(t_{2n-3})]\} \qquad (n = 2, 3, \cdots). \tag{3.3}$$

Since $\alpha'(t) < \frac{1}{10}$, it follows that $\beta'(t) < \frac{1}{10}$. Thus the equation above gives

$$t_{2n+1} - t_{2n-1} \leqq \tfrac{1}{2}\pi\{1 + \exp [\tfrac{1}{10}(t_{2n-1} - t_{2n-3})]\} \qquad (n = 2, 3, \cdots). \tag{3.4}$$

By (3.2), $t_3 = \frac{1}{2}\pi\{2 + \exp [\beta(\frac{1}{2}\pi)]\}$. Since $\beta(\frac{1}{2}\pi) < \frac{1}{10}(\frac{1}{2}\pi)$, it follows that

$$t_3 < \tfrac{1}{2}\pi(2 + e^{\pi/20}).$$

Since $t_1 = \frac{1}{2}\pi$,

$$t_3 - t_1 < \tfrac{1}{2}\pi(1 + e^{\pi/20}) < \tfrac{1}{2}\pi(4) = 2\pi.$$

If we use (3.4), it follows easily that if $t_{2n-1} - t_{2n-3} < 10$, then so is $t_{2n+1} - t_{2n-1}$. But $t_3 - t_1 < 2\pi < 10$. Thus by induction

$$t_{2n+1} - t_{2n-1} < 10 \qquad (n = 1, 2, \cdots).$$

Since, for large t, $\beta'(t) = o(1)$, clearly

$$\begin{aligned} \beta(t_{2n-1}) - \beta(t_{2n-3}) &= o(t_{2n-1} - t_{2n-3}) \\ &= o(10) = o(1). \end{aligned} \tag{3.5}$$

Thus (3.3) becomes

$$t_{2n+1} - t_{2n-1} = \pi + o(1).$$

This together with the fact that $t_{2n+1} - t_{2n} = \frac{1}{2}\pi$ completes the proof of the lemma.

Let $a_1 = \exp[-\beta(\frac{1}{2}\pi)]$ and let

$$\begin{aligned} a_{2n+1} &= \exp[-\beta(t_{2n+1}) + \beta(t_{2n-1})] \qquad (n = 1, 2, \cdots), \\ a_{2n} &= 1 \qquad (n = 0, 1, 2, \cdots). \end{aligned} \tag{3.6}$$

Let

$$\phi(t) = a_n^2, \qquad t_n \leqq t < t_{n+1} \qquad (n = 0, 1, 2, \cdots). \tag{3.7}$$

LEMMA 2. *For large t*

$$\int_0^t |\phi(t) - 1|\, dt \leqq \alpha(t). \tag{3.8}$$

Proof. By (3.2) and (3.6)

$$\begin{aligned} \int_0^{t_{2n+2}} |1 - \phi(t)|\, dt &= \frac{\pi}{2} \sum_{k=0}^{n} \frac{1 - a_{2k+1}^2}{a_{2k+1}} \\ &= \frac{\pi}{2} \frac{1 - a_1^2}{a_1} + \frac{\pi}{2} \sum_{k=1}^{n} \{\exp[\beta(t_{2k+1}) - \beta(t_{2k-1})] \\ &\qquad - \exp[-\beta(t_{2k+1}) + \beta(t_{2k-1})]\} \\ &= \frac{\pi}{2} \frac{1 - a_1^2}{a_1} + \pi \sum_{k=1}^{n} \sinh(\beta(t_{2k+1}) - \beta(t_{2k-1})). \end{aligned}$$

By (3.5) and the fact that $\sinh x = x + o(x)$, $x \to 0$, the above equation becomes

$$\int_0^{t_{2n+2}} |1 - \phi(t)|\, dt = \frac{\pi}{2} \frac{1 - a_1^2}{a_1} + [\pi + o(1)][\beta(t_{2n+1}) - \beta(t_1)].$$

For large n and $\delta > 0$ this becomes

$$\int_0^{t_{2n+2}} |1 - \phi(t)|\, dt \leqq \pi(1 + \delta)\beta(t_{2n+1}).$$

If we use Lemma 1, this becomes

$$\int_0^t |1 - \phi(t)| \, dt \leqq \pi(1 + \delta)\beta(t + \pi).$$

Since $\beta'(t) = o(1)$, this in turn gives

$$\int_0^t |1 - \phi(t)| \, dt \leqq \pi(1 + 2\delta)\beta(t).$$

If we use the definition of $\beta(t)$ and choose δ sufficiently small, (3.8) follows. Thus the lemma is proved.

Proof of Theorems II *and* IV. If A_n and b_n are constant, and if

$$X_n(t) = A_n \sin(a_n t + b_n), \tag{3.9}$$

then clearly $x(t) = X_n(t)$ satisfies the differential equation $x'' + \phi(t)x = 0$ over $t_n < t < t_{n+1}$, where $\phi(t)$ is defined as in (3.7). Thus in order that

$$x(t) = X_n(t), \qquad t_n \leqq t < t_{n+1} \qquad (n = 0, 1, 2, \cdots)$$

satisfy the differential equation over $(0, \infty)$, it is necessary and sufficient that $x(t)$ and $x'(t)$ be continuous. This in turn is equivalent to

$$X_{n-1}(t_n) = X_n(t_n), \qquad X'_{n-1}(t_n) = X'_n(t_n) \qquad (n = 1, 2, \cdots). \tag{3.10}$$

Writing (3.10) for n even and using (3.9) give

$$\begin{aligned} A_{2n-1} \sin(a_{2n-1}t_{2n} + b_{2n-1}) &= A_{2n} \sin(a_{2n}t_{2n} + b_{2n}), \\ a_{2n-1}A_{2n-1} \cos(a_{2n-1}t_{2n} + b_{2n-1}) &= a_{2n}A_{2n} \cos(a_{2n}t_{2n} + b_{2n}). \end{aligned} \tag{3.11}$$

Let

$$\begin{aligned} a_{2n-1}t_{2n} + b_{2n-1} &= \tfrac{1}{2}\pi \qquad (n = 1, 2, \cdots), \\ a_{2n}t_{2n} + b_{2n} &= \tfrac{1}{2}\pi \qquad (n = 0, 1, 2, \cdots). \end{aligned} \tag{3.12}$$

Then this defines b_n. Using (3.12) in (3.11) gives

$$A_{2n-1} = A_{2n} \qquad (n = 1, 2, \cdots). \tag{3.13}$$

For odd n (3.10) gives

$$\begin{aligned} A_{2n} \sin(a_{2n}t_{2n+1} + b_{2n}) &= A_{2n+1} \sin(a_{2n+1}t_{2n+1} + b_{2n+1}), \\ a_{2n}A_{2n} \cos(a_{2n}t_{2n+1} + b_{2n}) &= a_{2n+1}A_{2n+1} \cos(a_{2n+1}t_{2n+1} + b_{2n+1}). \end{aligned} \tag{3.14}$$

By (3.12),

$$a_{2n}t_{2n+1} + b_{2n} = a_{2n}(t_{2n+1} - t_{2n}) + \tfrac{1}{2}\pi.$$

From (3.2) and (3.6), it follows that

$$a_{2n}t_{2n+1} + b_{2n} = \pi.$$

Similarly

$$a_{2n+1}t_{2n+1} + b_{2n+1} = 0.$$

Using these last two equations, we see that (3.14) becomes

$$-a_{2n}A_{2n} = a_{2n+1}A_{2n+1} \qquad (n = 0, 1, 2, \cdots). \tag{3.15}$$

Let $A_0 = 1$. Then (3.13) and (3.15) define A_n for $n > 0$. Since (3.10) is satisfied, we have a solution of the differential equation over $(0, \infty)$. Since on the basis of the definitions of A_n, a_n, and b_n, $X_0(t) = \sin(t + \frac{1}{2}\pi)$, it follows that $x(0) = 1$, $x'(0) = 0$. From now on we denote this solution of the differential equation by $x_1(t)$.

Let us now replace (3.12) by

$$t_{2n} + b_{2n} = 0, \qquad a_{2n+1}t_{2n+1} + b_{2n+1} = \tfrac{1}{2}\pi \qquad (n = 0, 1, 2, \cdots), \tag{3.16}$$

and (3.13) and (3.15) by

$$\begin{aligned} A_{2n} &= A_{2n+1} && (n = 0, 1, 2, \cdots), \\ -a_{2n-1}A_{2n-1} &= a_{2n}A_{2n} && (n = 1, 2, \cdots). \end{aligned} \tag{3.17}$$

Then as above this gives rise to a solution of the differential equation which we shall denote by $x_2(t)$. By taking $A_0 = 1$ as before, it follows easily that $x_2(0) = 0$, $x_2'(0) = 1$.

We now consider the growth of $x_1(t)$. Clearly

$$\begin{aligned} x_1(t_{2n+2}) &= X_{2n+1}(t_{2n+2}) \\ &= A_{2n+1} \sin(a_{2n+1}t_{2n+2} + b_{2n+1}). \end{aligned}$$

By (3.12), $a_{2n+1}t_{2n+2} + b_{2n+1} = \frac{1}{2}\pi$. Thus

$$x_1(t_{2n+2}) = A_{2n+1}.$$

Clearly

$$\left|\frac{A_{2n+1}}{A_0}\right| = \left|\frac{A_1}{A_0}\frac{A_2}{A_1}\cdots\frac{A_{2n+1}}{A_{2n}}\right|.$$

If (3.13), (3.15) and the definition of A_n are used, this becomes

$$|A_{2n+1}| = \frac{1}{|a_1a_3\cdots a_{2n+1}|}.$$

From the definition of a_{2n+1}, (3.6), this becomes

$$|A_{2n+1}| = \exp[\beta(t_{2n+1}) - \beta(\tfrac{1}{2}\pi)].$$

Thus

$$|x_1(t_{2n+2})| = \exp[\beta(t_{2n+1}) - \beta(\tfrac{1}{2}\pi)].$$

But

$$\begin{aligned} \beta(t_{2n+1}) &\geqq \beta(t_{2n+2}) - \tfrac{1}{10}(t_{2n+2} - t_{2n+1}) \\ &\geqq \beta(t_{2n+2}) - \tfrac{1}{10}\pi. \end{aligned}$$

Also $\beta(\frac{1}{2}\pi) \leqq \frac{1}{20}\pi$. Thus

$$| x_1(t_{2n+2}) | \geqq \exp [\beta(t_{2n+2}) - \pi] \geqq \exp [(1 - \epsilon)\alpha(t_{2n+2})\pi^{-1}] e^{-\pi}.$$

Or

$$\limsup_{t\to\infty} \frac{\log | x_1(t) |}{\alpha(t)} \geqq \frac{1 - \epsilon}{\pi}. \tag{3.18}$$

Since ϵ is arbitrary, Theorem II follows.

In much the same way it follows that

$$\limsup_{t\to\infty} \frac{\log | x_2(t) |}{\alpha(t)} \leqq -\frac{1}{\pi},$$

and Theorem IV is proved.

4. Here we consider the case where the coefficient in the differential equation is a smooth function. Clearly there exists a continuous function $\phi_1(t)$ possessing any finite number of derivatives such that

$$\int_t^\infty | \phi_1(t) - \phi(t) | \, dt = O(e^{-t\alpha(t)}), \qquad t \to \infty. \tag{4.1}$$

Let us consider the differential equation

$$y'' + \phi_1(t)y = 0.$$

If x_1 is defined as in §3, then

$$x_1'' + \phi(t)x_1 = 0.$$

Thus, combining the two differential equations, we get

$$(y - x_1)'' + \phi(t)(y - x_1) = y[\phi(t) - \phi_1(t)]. \tag{4.2}$$

Consider the equation

$$x'' + \phi(t)x = f(t). \tag{4.3}$$

Then if x_1 and x_2 are defined as in §3,

$$x(t) = \int_t^\infty [x_1(t)x_2(\tau) - x_2(t)x_1(\tau)]f(\tau)\, d\tau$$

is a solution of (4.3). [This solution can be obtained by using the method of variation of constants, or it can be verified by putting it into the original equation.] Thus a solution of (4.2) is

$$y_1(t) - x_1(t) = \int^\infty [x_1(t)x_2(\tau) - x_2(t)x_1(\tau)]y_1(\tau)[\phi(\tau) - \phi_1(\tau)]\, d\tau. \tag{4.4}$$

By Theorem I

$$x_1(t), x_2(t), y_1(t), y_2(t) = O(e^{\alpha(t)}).$$

Thus (4.4) becomes

$$y_1(t) - x_1(t) = O\left(\int_t^\infty e^{\alpha(t)+\alpha(\tau)} e^{\alpha(\tau)} \mid \phi(\tau) - \phi_1(\tau) \mid d\tau\right)$$

$$= O\left(\int_t^\infty e^{3\alpha(\tau)} \mid \phi(\tau) - \phi_1(\tau) \mid d\tau\right).$$

Or integrating by parts and using (4.1), we get

$$y_1(t) - x_1(t) = O\left(e^{-(t-3)\alpha(t)} + \int_t^\infty e^{-(\tau-3)\alpha(\tau)}\, d\tau\right)$$

$$= O\left(e^{-(t-3)\alpha(t)} + \int_t^\infty e^{-(\tau-3)\alpha(t)}\, d\tau\right)$$

$$= O(e^{-(t-3)\alpha(t)}) = O(e^{-10\alpha(t)}).$$

But this and (3.18) imply that

$$\limsup_{t\to\infty} \frac{\log \mid y_1(t) \mid}{\alpha(t)} \geqq \frac{1}{\pi}.$$

Similarly it can be shown that there exists a $y_2(t)$ such that

$$y_2(t) - x_2(t) = O(e^{-10\alpha(t)})$$

and thus that

$$\limsup_{t\to\infty} \frac{\log \mid y_2(t) \mid}{\alpha(t)} \leqq -\frac{1}{\pi}.$$

Thus the differential equation

$$y'' - \phi_1(t)y = 0,$$

where $\phi_1(t)$ is a smooth function, has two solutions $y_1(t)$ and $y_2(t)$ which possess the extremal properties indicated in Theorems II and IV respectively.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

REGULAR CURVE-FAMILIES FILLING THE PLANE, II

By Wilfred Kaplan

Introduction. The present paper is a continuation of a previous one on the same subject.[1] The numbers of the theorems of I will be distinguished by the prefacing of the numeral I.

The problem to be considered here is the classification of regular curve-families filling the plane. Two families will be regarded as *equivalent* if one is image of the other under a homeomorphism of the plane onto itself. It is the enumeration of the classes of equivalent families which is our goal.

From I we have the results that the curves of each family F are open curves tending to infinity in both directions, that F can be pictured as a system with two order relations: $C_1 \mid C_2 \mid C_3$ and $\mid C_1, C_2, C_3 \mid^+$, termed a chordal system CS(F), and that, moreover, F forms a normal chordal system. This last structural feature of the family will serve as the basis of the classification.

We shall term F_1 *o-equivalent* to F_2 if F_1 is equivalent to F_2 under an orientation-preserving homeomorphism. The basic theorem of the classification can be stated as follows (see Part 3 below):

Theorem. *F_1 is o-equivalent to F_2 if and only if $CS(F_1)$ is isomorphic to $CS(F_2)$. For any F, $CS(F)$ is a normal chordal system and to every (abstract) normal chordal system E corresponds a curve-family F for which $CS(F)$ is isomorphic to E.*

By this theorem the "o-equivalence classes" of families can be enumerated, and an enumeration of the equivalence classes can thereby be obtained. This latter classification is made more effective by a *representation theorem*: *every normal chordal system can be represented by an isomorphic set K of non-intersecting chords on a circle.* The chordal relations in K are defined in the same way as in F. If we let K' denote the set of chords which is the image of K under reflection in a (fixed) diameter of the circle, then the final form of classification is thus: *F_1 is equivalent to F_2 if and only if K_1 is isomorphic either to K_2 or to K_2'.* Here K_i denotes the set of chords corresponding to F_i ($i = 1, 2$). The effectiveness of this classification lies in the fact that the isomorphism of two sets of chords is determined solely by the order of end-points on the circumference.

In Part 1 it will be established that to every abstract normal chordal system E corresponds a curve-family F which generates it. The normal subdivision of E into sets $\lambda(V_\alpha) = V_\alpha \cup \theta(V_\alpha)$, where $\theta(V_\alpha)$ is half-parallel, is used. The course

Received by the Editors of the Annals of Mathematics, February 21, 1940, accepted by them and later transferred to this Journal. The material in this paper and the preceding one was presented to the American Mathematical Society, September 5, 1939.

[1] See Kaplan, *Regular curve-families filling the plane*, I, this Journal, vol. 7(1940), pp. 154–185. A bibliography is given in that paper.

of procedure is thus: in §§1.1 and 1.2 the problem of how to verify that two normal chordal systems E and E' are isomorphic is studied, and that problem is reduced to the comparison of only two types of triples in each $\lambda(V_\alpha)$ and $\lambda(V'_\alpha)$; in §1.3 the plan of construction of the family F is described; in §§1.4 and 1.5 a curve-family is constructed to represent each $\lambda(V_\alpha)$ of E and in §1.6 the regularity of each such family is established; in §§1.7 and 1.8 these families are pieced together to yield a regular curve-family F filling the interior of a circle; in §1.9 the criteria of §§1.1 and 1.2 are applied to show that F is isomorphic to E.

In Part 2 it will be established that two curve-families F_1 and F_2 generate isomorphic chordal systems if and only if they are o-equivalent: §2.1 is introductory; §2.2 establishes a homeomorphism between the sets $\theta(V^1_\alpha)$ and $\theta(V^2_\alpha)$ by means of maps on parallel lines; in §§2.3 and 2.4 "isolating curves" $\gamma^1_{\alpha,k}$ and $\gamma^2_{\alpha,k}$ are found which cut off suitable neighborhoods $\omega^1_{\alpha,k}$ and $\omega^2_{\alpha,k}$ of the curves $C^1_{\alpha,k}$ and $C^2_{\alpha,k}$ in $\theta(V^1_\alpha)$ and $\theta(V^2_\alpha)$ respectively; in §2.5 the $\omega^1_{\alpha,k}$ and and $\omega^2_{\alpha,k}$ are adjusted in preparation for a homeomorphism; in §2.6 the homeomorphism of $\theta(V^1_\alpha)$ on $\theta(V^2_\alpha)$ is extended, by means of the sets $\omega^1_{\alpha,k}$ and $\omega^2_{\alpha,k}$, to the boundaries, yielding a map of $\lambda(V^1_\alpha)$ on $\lambda(V^2_\alpha)$; this is then extended to give an o-homeomorphism of F_1 on F_2, as desired.

In Part 3 these results will be used to give the classification as described above.

In the Appendix there will be a brief discussion of the application of the classification method to families with many singularities.

1. Regular curve-families corresponding to normal chordal systems

1.1. **Reduction of problem to case of $V \cup \theta(V)$.** In this section we shall establish the theorem that to every normal chordal system E corresponds a regular curve-family F filling the plane and such that $\mathrm{CS}(F)$ is isomorphic to E. The precise construction of the family F will be given below. However, in order to establish the isomorphism of $\mathrm{CS}(F)$ and E, it will be necessary to verify that for every triple C_1, C_2, C_3 in F the same chordal relation holds as for the corresponding triple in E. Our first step will be to anticipate this difficulty and to take advantage of the normal subdivision to show that the verification need be made only for a restricted class of triples. (See Theorems 1, 4 below.)

We shall let E be a fixed abstract normal chordal system, with normal subdivision by sets V_α of $E_0 = a_1 \cup \delta(a_1)$ and by sets V^*_α of $E^*_0 = a_1 \cup \delta^*(a_1)$ as in I, §3.3.

LEMMA 1. *If V is a non-void subset of E for which $\delta(a)$ can be determined as a single-valued function for all a in V in such a way that*

$$[a \cup \delta^*(a)]\cdot[a' \cup \delta^*(a')] = 0 \tag{1}$$

for every pair a, a' of distinct elements in V, then V is cyclic.

Proof. If $a' \mid a \mid a''$ for a, a', a'' in V, then $a'' \subset \delta(a)$, $a' \subset \delta^*(a)$ for proper naming of a'' and a'. This contradicts (1). Hence $\mid a', a, a'' \mid^{\pm}$ for every triple in V and V is cyclic.

LEMMA 2. *Let $V_1, V_2, \cdots, V_p$ be cyclic subsets of E with $\theta(V_i)$ determined uniquely $(i = 1, 2, \cdots, p)$ $(p \geqq 2)$. If $\lambda(V_i)$ and $\lambda(V_{i+1})$ are adjacent with common element k_i for $i = 1, 2, \cdots, p - 1$, and if the k_i are distinct, then the set $V = \sum_{i=1}^{p} V_i - \sum_{i=1}^{p-1} k_i$ is either void, in which case $\sum_1^p \lambda(V_i) = E$, or else V is cyclic and*

$$\theta(V) = \sum_1^p \theta(V_i) \cup \sum_1^{p-1} k_i, \qquad \lambda(V) = \sum_1^p \lambda(V_i).$$

Proof. Suppose that $p = 2$ and that V is not void. Let $V_1 = k_1 \cup \sum_j a_j$, $V_2 = k_1 \cup \sum_l b_l$, with $\theta(V_1) = \delta(k_1) \cdot \prod_j \delta(a_j)$ and $\theta(V_2) = \delta^*(k_1) \cdot \prod_l \delta(b_l)$. By Theorem I.27

$$(1) \qquad [a_j \cup \delta^*(a_j)] \cdot [a_{j'} \cup \delta^*(a_{j'})] = 0, \qquad j \neq j',$$

$$(2) \qquad [b_l \cup \delta^*(b_l)] \cdot [b_{l'} \cup \delta^*(b_{l'})] = 0, \qquad l \neq l',$$

$$(3) \qquad a_j \cup \delta^*(a_j) \subset \delta(k_1),$$

$$(4) \qquad [b_l \cup \delta^*(b_l)] \cdot [\delta(k_1) \cup k_1] = 0.$$

From (3) and (4) follows

$$(5) \qquad [a_j \cup \delta^*(a_j)] \cdot [b_l \cup \delta^*(b_l)] = 0.$$

(1), (2) and (5) imply by Lemma 1 that $(V_1 \cup V_2) - k_1$ is cyclic and

$$\begin{aligned} \theta((V_1 \cup V_2) - k_1) &= \prod \delta(a_j) \cdot \prod \delta(b_l) \\ &= \prod \delta(a_j) \cdot \prod \delta(b_l) \cdot [\delta(k_1) \cup \delta^*(k_1) \cup k_1] \\ &= [\delta(k_1) \cdot \prod \delta(a_j)] \cup [\delta^*(k_1) \cdot \prod \delta(b_l)] \cup k_1 \\ &= \theta(V_1) \cup \theta(V_2) \cup k_1 \end{aligned}$$

unless $(V_1 \cup V_2) - k_1$ has only one element a, e.g., in V_1. In that case

$$\theta((V_1 \cup V_2) - k_1) = \theta(a) = \delta(a) = (\delta(a) \cdot \delta(k_1)) \cup \delta^*(k_1) \cup k_1 = \theta(V_1) \cup \theta(V_2) \cup k_1$$

if we define $\theta(a)$ as $\delta(a)$.

Finally we have

$$\lambda(V) = \sum a_j \cup \sum b_l \cup \theta(V_1) \cup \theta(V_2) \cup k_1 = \lambda(V_1) \cup \lambda(V_2).$$

Thus the theorem holds when V is non-void and $p = 2$. If $p = 2$ but V is void, then $V_1 = V_2 = k_1$, $\lambda(V_1) = k_1 \cup \delta(k_1)$, $\lambda(V_2) = k_1 \cup \delta^*(k_1)$ and $\lambda(V_1) \cup \lambda(V_2) = E$ by Theorem I.26. Thus, in either case, the theorem holds for $p = 2$.

Now suppose the theorem established for $p = r$. Then let $p = r + 1$. Let $W = \sum_1^r V_i - \sum_1^{r-1} k_i$ and suppose that W is not void. We then have

$$\theta(W) = \sum_1^r \theta(V_i) \cup \sum_1^{r-1} k_i .$$

Further, since the $\lambda(V_i)$ are adjacent,

$$\theta(V_1) \subset \delta(k_1),\ \theta(V_2) \subset \delta^*(k_1) \subset \delta(k_2),\ \cdots,$$
$$\theta(V_r) \subset \delta^*(k_{r-1}) \subset \delta(k_r),\ \theta(V_{r+1}) \subset \delta^*(k_r)$$

for proper choices of $\delta(k_i)$. We hence have

$$\delta(k_1) \subset \delta(k_2) \subset \cdots \subset \delta(k_r);\ k_1 \subset \delta(k_2),\ k_2 \subset \delta(k_3),\ \cdots,\ k_{r-1} \subset \delta(k_r),$$

whence $\theta(W) \subset \delta(k_r)$, while $\theta(V_{r+1}) \subset \delta^*(k_r)$. Thus $\lambda(W)$ and $\lambda(V_{r+1})$ are adjacent. Applying the theorem for the case $p = 2$, we obtain that the set $V = (W \cup V_{r+1}) - k_r$ is either void, in which case

$$E = \lambda(W) \cup \lambda(V_{r+1}) = \sum_1^r \lambda(V_i) \cup \lambda(V_{r+1}) = \sum_1^{r+1} \lambda(V_i),$$

or else cyclic and

$$\lambda(V) = \lambda(W) \cup \lambda(V_{r+1}) = \sum_1^{r+1} \lambda(V_i),$$

$$\theta(V) = \theta(W) \cup \theta(V_{r+1}) \cup k_r = \sum_1^{r+1} \theta(V_i) \cup \sum_1^r k_i .$$

Hence the theorem holds for $p = r + 1$ when W is not void. If W were void, then $V_r = k_{r-1}$, whence $k_r = k_{r-1}$, contrary to the assumption that the k_i are distinct. Therefore this case cannot arise. By induction we now conclude that the theorem holds for all p.

LEMMA 3. *Under the assumptions of Lemma 2, the chordal relations for each triple a, b, c in the set $\sum_1^p \lambda(V_i)$ are determined by the chordal relations of triples in each of $\lambda(V_1)$, $\lambda(V_2)$, $\cdots$, $\lambda(V_p)$.*

Proof. Suppose $p = 2$. Then the only case requiring discussion is when, for example, a and b are in $\lambda(V_1)$, c is in $\lambda(V_2)$ and $c \neq k_1$. If $a = k_1$, then $b \subset \delta(k_1)$, $c \subset \delta^*(k_1)$, whence $b \mid k_1 \mid c$ and $b \mid a \mid c$. If $a \neq k_1$, $b \neq k_1$, then we have both (1) $b \mid k_1 \mid c$ and (2) $a \mid k_1 \mid c$. By Axiom 3.4 $a \mid k_1 \mid b$ is not true. If (3) $k_1 \mid a \mid b$, then (1) and (3) give, by Theorem I.17, $a \mid b \mid c$. If (4) $k_1 \mid b \mid a$, then (2) and (4) give, by Theorem I.17, $c \mid b \mid a$. If (5) $\mid k_1, a, b \mid^{\pm}$, then (1) and (5) give, by Axiom 3.2, $[a, c, b] \sim [k_1, a, b]$. Hence in all cases the relation for a, b, c is determined by that for a, k_1, b in $\lambda(V_1)$. Thus the theorem holds for $p = 2$.

Suppose now that the theorem has been established for $p = r$, and set $W = \sum_1^r V_i - \sum_1^{r-1} k_i$ as above. Then, since $\lambda(W)$ is adjacent to $\lambda(V_{r+1})$, the chordal relations in $\lambda(V) = \lambda(W) \cup \lambda(V_{r+1})$ are determined by those in $\lambda(W)$ and $\lambda(V_{r+1})$ individually, hence in $\lambda(V_1)$, $\lambda(V_2)$, $\cdots$, $\lambda(V_{r+1})$ individually. Thus the theorem holds for $p = r + 1$ and, by induction, for all p.

LEMMA 4. *The chordal relations in E_0 are determined by those in the sets $\lambda(V_\alpha)$ individually.*

Proof. Let a_1, a_2, a_3 be three distinct elements of E_0, not all in the same $\lambda(V_\alpha)$. Let $a_i \subset \lambda(V_{\alpha_i})$ $(i = 1, 2, 3)$. Let α_0 be the largest sequence such that we have

$$\alpha_1 = \alpha_0, \lambda_1, \lambda_2, \cdots, \lambda_r, \qquad r \geqq 0,$$

$$\alpha_2 = \alpha_0, \mu_1, \mu_2, \cdots, \mu_s, \qquad s \geqq 0,$$

$$\alpha_3 = \alpha_0, \nu_1, \nu_2, \cdots, \nu_t, \qquad t \geqq 0.$$

Suppose first that $r > 0$, $s > 0$, $t > 0$. Then the possible situations are exemplified by either $\lambda_1 \neq \mu_1$, $\mu_1 = \nu_1$ or $\lambda_1 \neq \mu_1$, $\mu_1 \neq \nu_1$, $\lambda_1 \neq \nu_1$. In each of these two cases, let α_4 be the largest sequence such that

$$\alpha_2 = \alpha_4, m_1, m_2, \cdots, m_\sigma, \qquad \sigma \geqq 0,$$

$$\alpha_3 = \alpha_4, n_1, n_2, \cdots, n_\tau, \qquad \tau \geqq 0,$$

$$\alpha_4 = \alpha_0, \mu_1, \mu_2, \cdots, \mu_u, \qquad u \geqq 0.$$

Then, by the normality of E, the sets

$$W_1 = V_{\alpha_4, m_1, m_2, \cdots, m_\sigma} = V_{\alpha_2},$$

$$W_2 = V_{\alpha_4, m_1, m_2, \cdots, m_{\sigma-1}},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots,$$

$$W_\sigma = V_{\alpha_4, m_1},$$

$$W_{\sigma+1} = V_{\alpha_4},$$

$$W_{\sigma+2} = V_{\alpha_4, n_1},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots,$$

$$W_{\sigma+\tau+1} = V_{\alpha_4, n_1, n_2, \cdots, n_\tau} = V_{\alpha_3}$$

satisfy the conditions of Lemma 3. Hence the chordal relations in $\lambda(X_1) = \sum_1^{\sigma+\tau+1} \lambda(W_p)$ are determined by those in a certain set of $\lambda(V_\alpha)$ individually.

Similarly the sets X_1 and

$$\begin{aligned} X_2 &= V_{\alpha_0,\mu_1,\mu_2,\cdots,\mu_{u-1}}, \\ X_3 &= V_{\alpha_0,\mu_1,\mu_2,\cdots,\mu_{u-2}}, \\ &\cdots\cdots\cdots\cdots, \\ X_u &= V_{\alpha_0,\mu_1}, \\ X_{u+1} &= V_{\alpha_0}, \\ X_{u+2} &= V_{\alpha_0,\lambda_1}, \\ &\cdots\cdots\cdots\cdots, \\ X_{u+r+1} &= V_{\alpha_0,\lambda_1,\lambda_2,\cdots,\lambda_r} = V_{\alpha_1} \end{aligned}$$

satisfy the conditions of Lemma 3. Hence the chordal relations in $\lambda(Y) = \sum_1^{u+r+1} \lambda(X_q)$ are determined by those in sets $\lambda(X_1), \lambda(X_2), \cdots, \lambda(X_{u+r+1})$ individually, hence in the sets $\lambda(V_\alpha)$ individually. Since $\lambda(Y)$ includes $\sum_{i=1}^{3} \lambda(V_{\alpha_i})$ and hence a_1, a_2, a_3, the theorem is established for $r > 0$, $s > 0$, $t > 0$. If one of these indices $= 0$, we deduce the same result by a simplification of the above method.

THEOREM 1. *Let E be the normal chordal system considered above. Let E' be a second normal chordal system, subdivided normally by sets V'_α for $\alpha \subset A$, and by sets V'^*_α for α in A^*, with*

$$\sum_\alpha \lambda(V'_\alpha) = E'_0 = a'_1 \cup \delta(a'_1), \qquad \sum_\alpha \lambda(V'^*_\alpha) = E'^*_0 = a'_1 \cup \delta^*(a'_1).$$

Let $f(a) = a'$ be a one-to-one transformation of E on E' such that f is an isomorphism of each $\lambda(V_\alpha)$ on $\lambda(V'_\alpha)$, of each $\lambda(V^*_\alpha)$ on $\lambda(V'^*_\alpha)$. Then E is isomorphic to E'.

Proof. Since $f(\lambda(V_\alpha)) = \lambda(V'_\alpha)$ for all α, necessarily $f(c_\alpha) = c'_\alpha$ and similarly $f(c^*_\alpha) = c'^*_\alpha$. By Lemma 4 the chordal relations in E are determined by those of the $\lambda(V_\alpha)$ individually (and the way in which they are adjacent). Similarly those in E^*_0 are determined by those of the $\lambda(V^*_\alpha)$. E_0 and E^*_0 are both sets $\lambda(V)$ to which Lemma 3 can be applied. Hence the chordal relations in $E = E_0 \cup E^*_0$ are determined by those in E_0 and E^*_0, hence ultimately by those in the $\lambda(V_\alpha)$ and $\lambda(V^*_\alpha)$. The same condition holds for E' relative to the V'_α and V'^*_α. But the ranges of α are the same in both E_0 and E'_0 and in E^*_0 and E'^*_0, and f is an isomorphism on each $\lambda(V_\alpha)$ and $\lambda(V^*_\alpha)$ with $f(c_\alpha) = c'_\alpha$, $f(c^*_\alpha) = c'^*_\alpha$. Hence the chordal relations for triples in E are the same as those for the corresponding triples in E'. Thus E is isomorphic to E'. This proves the theorem.

We shall apply this theorem below after the curve-family F corresponding to E has been constructed. By means of it we are able to determine, only by comparing the sets $\lambda(V_\alpha)$ with the corresponding sets $\lambda(V'_\alpha)$ in $E' = \mathrm{CS}(F)$, that E is isomorphic to E'.

1.2. Simplification of the problem in a $\lambda(V_\alpha)$. We express $\lambda(V_\alpha)$ as $c_\alpha \cup \sum c_{\alpha,k} \cup \sum d_t^\alpha$ (with $d_0^\alpha = c_\alpha$) as in I, §4.1.

Notation. V_α^+ will denote the set of all elements $c_{\alpha,k}$ such that $| c_\alpha , c_{\alpha,k} , d_t^\alpha |^+$ for some d_t^α ; V_α^- will denote the set of all $c_{\alpha,k}$ such that $| c_\alpha , c_{\alpha,k} , d_t^\alpha |^-$ for some d_t^α .

THEOREM 2. $V_\alpha^+ \cdot V_\alpha^- = 0$. $V_\alpha^+ \cup V_\alpha^- = V_\alpha - c_\alpha$.

Proof. If $| c_\alpha , c_{\alpha,k} , d_{t_1}^\alpha |^\pm$ and $| c_\alpha , c_{\alpha,k} , d_{t_2}^\alpha |^\pm$, with $t_1 < t_2$, then $c_\alpha \mid d_{t_1}^\alpha \mid d_{t_2}^\alpha$, whence by Axiom 3.2 $[c_\alpha , c_{\alpha,k} , d_{t_1}^\alpha] \sim [c_\alpha , c_{\alpha,k} , d_{t_2}^\alpha]$. Hence $V_\alpha^+ \cdot V_\alpha^- = 0$. The second equation follows from condition (6) of normality. The theorem thus follows.

THEOREM 3. *The definition $c_{\alpha,k} < c_{\alpha,m}$ for $| c_\alpha , c_{\alpha,k} , c_{\alpha,m} |^+$ introduces a simple ordering in V_α^+ . The definition $c_{\alpha,k} < c_{\alpha,m}$ for $| c_\alpha , c_{\alpha,k} , c_{\alpha,m} |^-$ introduces a simple ordering in V_α^- . Further $c_{\alpha,k} < c_{\alpha,l} < c_{\alpha,m}$ implies $| c_{\alpha,k} , c_{\alpha,l} , c_{\alpha,m} |^+$ for $c_{\alpha,k} , c_{\alpha,l} , c_{\alpha,m}$ in V_α^+ , implies $| c_{\alpha,k} , c_{\alpha,l} , c_{\alpha,m} |^-$ for $c_{\alpha,k} , c_{\alpha,l} , c_{\alpha,m}$ in V_α^- .*

Proof. Take $c_{\alpha,k}$, $c_{\alpha,m}$ in V_α^+ . Then $| c_\alpha , c_{\alpha,k} , c_{\alpha,m} |^\pm$. Hence either $c_{\alpha,k} < c_{\alpha,m}$ or $c_{\alpha,m} < c_{\alpha,k}$. If $c_{\alpha,k} < c_{\alpha,l}$, $c_{\alpha,l} < c_{\alpha,m}$, then $| c_\alpha , c_{\alpha,k} , c_{\alpha,l} |^+$, $| c_\alpha , c_{\alpha,l} , c_{\alpha,m} |^+$, whence, by Axiom 3.1, $| c_\alpha , c_{\alpha,k} , c_{\alpha,m} |^+$, i.e., $c_{\alpha,k} < c_{\alpha,m}$. Further, by Axiom 3.1, $| c_{\alpha,k} , c_{\alpha,l} , c_{\alpha,m} |^+$. The case of V_α^- is treated in the same way.

We now renumber the $c_{\alpha,k}$ by allowing k to take negative values in such a way that the elements of V_α^+ are the $c_{\alpha,k}$ for $k = 1, 2, 3, \cdots$, of V_α^- are the $c_{\alpha,k}$ for $k = -1, -2, -3, \cdots$. The range of k may of course be finite, infinite or void in each case.

The possible types of triples in $\lambda(V_\alpha)$ can be then indicated as follows: (1) $d_{t_1}^\alpha$, $d_{t_2}^\alpha$, $d_{t_3}^\alpha$; (2) $c_{\alpha,k}$, $c_{\alpha,l}$, $c_{\alpha,m}$ with (a) k, l and $m > 0$ or (b) k, l and $m < 0$; (3) $c_{\alpha,k}$, $c_{\alpha,l}$, $c_{\alpha,m}$, with $k > 0$, $l < 0$ and (a) $m > 0$ or (b) $m < 0$; (4) $d_{t_1}^\alpha$, $c_{\alpha,k}$, $c_{\alpha,l}$, with $k > 0$, $l < 0$; (5) $d_{t_1}^\alpha$, $c_{\alpha,k}$ and $c_{\alpha,l}$ with (a) k and $l > 0$ or (b) k and $l < 0$; (6) $d_{t_1}^\alpha$, $d_{t_2}^\alpha$ and $c_{\alpha,k}$ with (a) $k > 0$ or (b) $k < 0$.

From the given structure of $\lambda(V_\alpha)$ certain restrictions are put on these relations, as follows: the relations of type (1) are determined by the relative sizes of t_1 , t_2 , t_3 ; for the relations of type (2) and (3) necessarily $| c_{\alpha,k} , c_{\alpha,l} , c_{\alpha,m} |^\pm$; for those of types (4) and (5) with $t_1 = 0$ $| d_0^\alpha , c_{\alpha,k} , c_{\alpha,l} |^\pm$, with $t_1 > 0$ $| d_{t_1}^\alpha , c_{\alpha,k} , c_{\alpha,l} |^\pm$ or $c_{\alpha,k} \mid d_{t_1}^\alpha \mid c_{\alpha,l}$ (by Theorem I.28).

LEMMA. *If, in addition to the restrictions on chordal relations in $\lambda(V_\alpha)$ implicit in its definition, the following relations are known: (α) all those for triples c_α , $c_{\alpha,k}$, d_t^α and (β) all those for triples $c_{\alpha,k}$, $c_{\alpha,l}$, $c_{\alpha,m}$ such that k, l, and m have the same sign and such that $B(c_{\alpha,k}) \equiv B(c_{\alpha,l}) \equiv B(c_{\alpha,m})$, then the remaining relations in $\lambda(V_\alpha)$ are uniquely determined.*

Proof. From (α) it follows that for each $c_{\alpha,k}$ it is known whether $k > 0$ or $k < 0$. Further the sets $B(c_{\alpha,k})$ and the element $d_{t_{\alpha,k}}^\alpha$ are thereby given.

The relations of *type* (1) are implicit in the definition of $\lambda(V_\alpha)$.

Consider a triple of *type* (2)(a). If $B(c_{\alpha,k}) \equiv B(c_{\alpha,l}) \equiv B(c_{\alpha,m})$, then by ($\beta$)

the relation is known. Assume then that this condition does not hold. It follows from Theorem 3 that, when we have determined the order relations in V_α^+, then all relations in V_α^+ are known. We can therefore restrict attention to a pair such as $c_{\alpha,k}$, $c_{\alpha,l}$ in V_α^+. There are the following cases: (i) $t_{\alpha,k} < t_{\alpha,l}$; (ii) $t_{\alpha,k} = t_{\alpha,l} = t^*$, with $d_{t^*}^\alpha \subset B(c_{\alpha,k})$ and $d_{t^*}^\alpha \not\subset B(c_{\alpha,l})$. In case (i) choose t_0 so that $t_{\alpha,k} < t_0 < t_{\alpha,l}$. Then $| c_\alpha, c_{\alpha,k}, d_{t_0}^\alpha |^+$ and $c_\alpha | d_{t_0}^\alpha | c_{\alpha,l}$, from Theorem 3 and Theorem I.28 respectively. Hence, by Axiom 3.2, $| c_\alpha, c_{\alpha,k}, c_{\alpha,l} |^+$ and $c_{\alpha,k} < c_{\alpha,l}$. In case (ii) the same reasoning applies with t^* replacing t_0. Hence all relations of type (2)(a) are determined, and similarly for those of type (2)(b).

Consider a triple of *type* (3)(a). From the result of the preceding paragraph we are given whether $c_{\alpha,k} < c_{\alpha,m}$ or $c_{\alpha,m} < c_{\alpha,k}$. Suppose (without restricting generality) the first case. Then (1) $| c_\alpha, c_{\alpha,k}, c_{\alpha,m} |^+$. Choose $t_0 >$ max $(t_{\alpha,m}, t_{\alpha,l})$. Then $| c_\alpha, c_{\alpha,m}, d_{t_0}^\alpha |^+$ and $| c_\alpha, c_{\alpha,l}, d_{t_0}^\alpha |^-$. Hence, by Axiom 3.1, $| c_\alpha, c_{\alpha,m}, c_{\alpha,l} |^+$. This and (1) give, by Axiom 3.1, $| c_{\alpha,k}, c_{\alpha,m}, c_{\alpha,l} |^+$. Hence in the case (3)(a) the relation is determined, and similarly in case (3)(b).

Consider a triple of *type* (4). There are the following cases: (i) $d_{t_1}^\alpha \subset B(c_{\alpha,k})$, $\not\subset B(c_{\alpha,l})$; (ii) $d_{t_1}^\alpha \not\subset B(c_{\alpha,k})$, $\subset B(c_{\alpha,l})$; (iii) $d_{t_1}^\alpha \subset B(c_{\alpha,k})$, $\subset B(c_{\alpha,l})$; (iv) $d_{t_1}^\alpha \not\subset B(c_{\alpha,k})$, $\not\subset B(c_{\alpha,l})$. In case (i) we have $t_1 > 0$ and $| c_\alpha, c_{\alpha,k}, d_{t_1}^\alpha |^+$, $c_\alpha | d_{t_1}^\alpha | c_{\alpha,l}$, whence, by Axiom 3.2, $c_{\alpha,k} | d_{t_1}^\alpha | c_{\alpha,l}$. Case (ii) is treated in the same way. In case (iii) $| c_\alpha, c_{\alpha,k}, d_{t_1}^\alpha |^+$, $| c_\alpha, c_{\alpha,l}, d_{t_1}^\alpha |^-$, whence, by Axiom 3.1, $| c_{\alpha,k}, d_{t_1}^\alpha, c_{\alpha,l} |^+$. In case (iv), if $t_1 = 0$, then take $t_0 >$ max $(t_{\alpha,k}, t_{\alpha,l})$. Hence $| d_0^\alpha, c_{\alpha,k}, d_{t_0}^\alpha |^+$, $| d_0^\alpha, c_{\alpha,l}, d_{t_0}^\alpha |^-$, whence, by Axiom 3.1, $| c_{\alpha,l}, d_0^\alpha, c_{\alpha,k} |^+$. If $t_1 > 0$, then $c_\alpha | d_{t_1}^\alpha | c_{\alpha,k}$, $c_\alpha | d_{t_1}^\alpha | c_{\alpha,l}$. We know that $c_{\alpha,k} | d_{t_1}^\alpha | c_{\alpha,l}$ or $| c_{\alpha,k}, d_{t_1}^\alpha, c_{\alpha,l} |^\pm$. But $c_{\alpha,k} | d_{t_1}^\alpha | c_{\alpha,l}$ contradicts Axiom 3.4. Hence $| c_{\alpha,k}, d_{t_1}^\alpha, c_{\alpha,l} |^\pm$. Thus, by Axiom 3.2, $[c_{\alpha,k}, d_{t_1}^\alpha, c_{\alpha,l}] \sim [c_{\alpha,k}, c_\alpha, c_{\alpha,l}] \sim [c_{\alpha,k}, d_0^\alpha, c_{\alpha,l}]$. Hence, by the case $t_1 = 0$ above, $| c_{\alpha,k}, d_{t_1}^\alpha, c_{\alpha,l} |^+$. Hence in case (4) the relations are determined.

In *case* (5)(a) we can assume $c_{\alpha,k} < c_{\alpha,l}$, whence $| c_\alpha, c_{\alpha,k}, c_{\alpha,l} |^+$. This covers the case $t_1 = 0$. Suppose $t_1 > 0$. There are then three cases: (i) $d_{t_1}^\alpha \subset B(c_{\alpha,l})$; (ii) $d_{t_1}^\alpha \subset B(c_{\alpha,k})$ and $\not\subset B(c_{\alpha,l})$; (iii) $d_{t_1}^\alpha \not\subset B(c_{\alpha,k})$. In case (i) we have $| c_\alpha, c_{\alpha,l}, d_{t_1}^\alpha |^+$, whence, by Axiom 3.1, $| c_{\alpha,k}, c_{\alpha,l}, d_{t_1}^\alpha |^+$. In case (ii) we have $| c_\alpha, c_{\alpha,k}, d_{t_1}^\alpha |^+$ and $c_\alpha | d_{t_1}^\alpha | c_{\alpha,l}$, whence, by Axiom 3.2, $c_{\alpha,k} | d_{t_1}^\alpha | c_{\alpha,l}$. In case (iii) we have $c_\alpha | d_{t_1}^\alpha | c_{\alpha,k}$, $c_\alpha | d_{t_1}^\alpha | c_{\alpha,l}$, whence, by Axiom 3.4, $c_{\alpha,k} | d_{t_1}^\alpha | c_{\alpha,l}$ is not true. Hence $| c_{\alpha,k}, d_{t_1}^\alpha, c_{\alpha,l} |^\pm$ and, by Axiom 3.2, $[c_{\alpha,k}, d_{t_1}^\alpha, c_{\alpha,l}] \sim [c_{\alpha,k}, c_\alpha, c_{\alpha,l}]$. Hence $| c_{\alpha,k}, d_{t_1}^\alpha, c_{\alpha,l} |^-$. This covers the case (5)(a). Case (5)(b) is treated similarly.

Consider a triple of *type* (6)(a). We can assume $0 < t_1 < t_2$, the case $t_1 = 0$ being covered by the information (α). We have then $c_\alpha | d_{t_1}^\alpha | d_{t_2}^\alpha$. There are (by Theorem I.40) the following cases: (i) $d_{t_1}^\alpha \subset B(c_{\alpha,k})$; (ii) $d_{t_1}^\alpha \not\subset B(c_{\alpha,k})$, $d_{t_2}^\alpha \subset B(c_{\alpha,k})$; (iii) $d_{t_2}^\alpha \not\subset B(c_{\alpha,k})$. In case (i) we have $| c_\alpha, c_{\alpha,k}, d_{t_1}^\alpha |^+$, whence, by Axiom 3.2, $c_{\alpha,k} | d_{t_1}^\alpha | d_{t_2}^\alpha$. In case (ii) $c_\alpha | d_{t_1}^\alpha | c_{\alpha,k}$ and $| c_\alpha, c_{\alpha,k}, d_{t_2}^\alpha |^+$. $c_{\alpha,k} | d_{t_1}^\alpha | d_{t_2}^\alpha$ contradicts Axiom 3.4. $d_{t_1}^\alpha | c_{\alpha,k} | d_{t_2}^\alpha$ would imply $c_\alpha | c_{\alpha,k} | d_{t_2}^\alpha$, by Axiom 3.3, and this is a contradiction. $d_{t_1}^\alpha | d_{t_2}^\alpha | c_{\alpha,k}$ would imply $c_\alpha | d_{t_2}^\alpha | c_{\alpha,k}$, and this also is a contradiction. Hence $| c_{\alpha,k}, d_{t_1}^\alpha, d_{t_2}^\alpha |^\pm$, whence, by Axiom

3.2, $[c_{\alpha,k}, d^\alpha_{t_1}, d^\alpha_{t_2}] \sim [c_{\alpha,k}, c_\alpha, d^\alpha_{t_2}]$ and $| c_{\alpha,k}, d^\alpha_{t_1}, d^\alpha_{t_2} |^-$. In case (iii), $c_\alpha \mid d^\alpha_{t_2} \mid c_{\alpha,k}$, whence, by Theorem I.17, $d^\alpha_{t_1} \mid d^\alpha_{t_2} \mid c_{\alpha,k}$. This covers case (6)(a). Case (b) is covered similarly. The lemma is now established.

THEOREM 4. *Let $\lambda(V_\alpha) = c_\alpha \cup \sum c_{\alpha,k} \cup \sum d^\alpha_t$ and $\lambda(V'_\alpha)$ be subsets in the normal subdivisions of chordal systems E and E' respectively. Let $f(a) = a'$ be a one-to-one transformation of $\lambda(V_\alpha)$ on $\lambda(V'_\alpha)$ with $f(c_\alpha) = c'_\alpha$, $f(c_{\alpha,k}) = c'_{\alpha,k}$, $f(d^\alpha_t) = d'^\alpha_t$. If further $[a_1, a_2, a_3] \sim [f(a_1), f(a_2), f(a_3)]$ for each triple a_1, a_2, a_3 of types (α) and (β), then f maps $\lambda(V_\alpha)$ isomorphically on $\lambda(V'_\alpha)$.*

This theorem, together with Theorem 1, will be used to simplify the verification of the isomorphism between E and the curve-family to be constructed below.

1.3. Intuitive outline of procedure.

We shall in the following sections construct a regular curve-family F corresponding to the above abstract normal chordal system E. However, instead of taking the family as filling the plane, we shall take it as filling the interior of the circle and in such a way that each curve joins a pair of points on the circumference, no two curves having a common limit point.

If we consider a set V_α in E, we see that its structure is that of a set of chords in the circle such that no one chord separates any two others of the set. Such a set, with the addition of certain points on the circumference, forms a simple closed curve G. The elements of $\theta(V_\alpha)$, regarded as curves, would then have to lie in the interior of the region bounded by G. Their limit points will lie on the gaps in G left by removing the chords of V_α.

We can make the situation simpler by assuming that the curves of $\theta(V_\alpha)$ and the chord c_α all cross just once the radius perpendicular to c_α. If we let P be a point varying along that radius and $\varphi_1(P)$ and $\varphi_2(P)$ be the angular coördinates of the limit points of the curve through P, then, as P moves from the mid-point of c_α towards the circumference, $\varphi_1(P)$ and $\varphi_2(P)$ change monotonely. $\varphi_1(P)$ and $\varphi_2(P)$ have discontinuities when one chord or several chords of V_α must be skipped over.

Our first step in the construction will be then to determine two monotone functions φ_1 and φ_2 whose discontinuities correspond to the gaps determined by the chords of V_α. These functions will place the limit points of the curves of $\theta(V_\alpha)$ and leave gaps in which we can place the chords of V_α. The joining of the limit points of the curves of $\theta(V_\alpha)$ in such a way as to get curves of a regular curve-family is then established by a limiting process.

In this way a family corresponding to $\lambda(V_1)$ can be set up, with a bounding diameter as the element C_1. On the other side of this diameter we construct $\lambda(V_1^*)$ in the same way. Then in the segments of the circle determined by the chords of V_1 we fit in the $\lambda(V_{1,k})$. Proceeding indefinitely in this way we fill out the interior of the circle with a regular curve-family of the desired structure. (See Figure 1.)

1.4. **Preliminary construction of curves corresponding to** $c_\alpha \cup \theta(V_\alpha)$.

Notation. Let $0 \leqq t_0 < \infty$. $Y_\alpha^+(t_0)$ (or $Y_\alpha^-(t_0)$) will then denote the set of all $c_{\alpha,k}$ in V_α^+ (or V_α^-) for which $t_{\alpha,k} = t_0$ and $d_{t_0}^\alpha \subset B(c_{\alpha,k})$. $Z_\alpha^+(t_0)$ (or $Z_\alpha^-(t_0)$) will denote the set of all $c_{\alpha,k}$ in V_α^+ (or V_α^-) for which $t_{\alpha,k} = t_0$ and $d_{t_0}^\alpha \not\subset B(c_{\alpha,k})$.

Definition. If $Y_\alpha^+(t_0) \cup Z_\alpha^+(t_0)$ (or $Y_\alpha^-(t_0) \cup Z_\alpha^-(t_0)$) is non-void, then $d_{t_0}^\alpha$ is a *discontinuity* with respect to V_α^+ (or V_α^-). If $Y_\alpha^+(t_0) \neq 0$ and $Z_\alpha^+(t_0) = 0$ (or

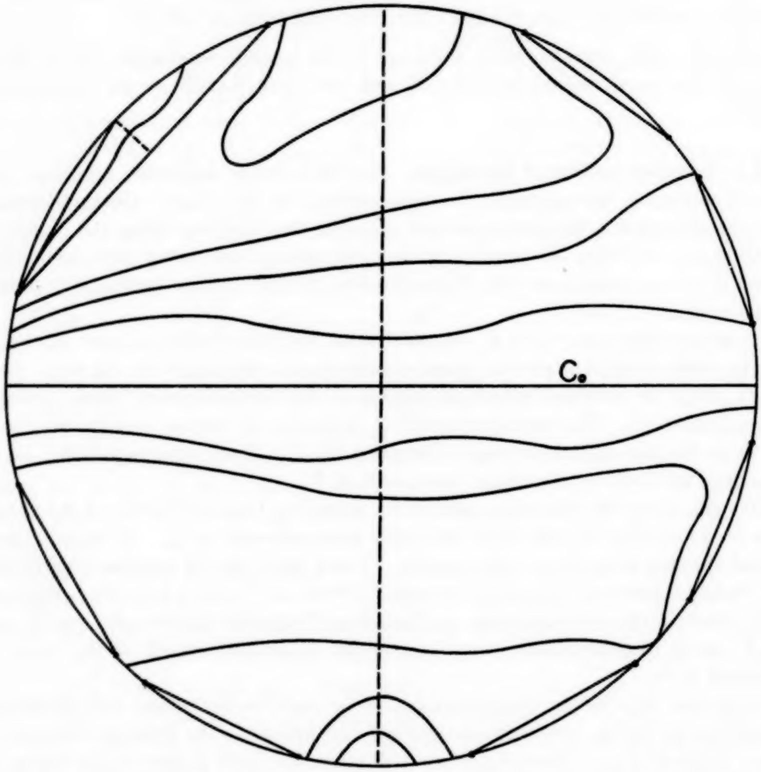

Fig. 1. Representation of an Abstract Chordal System by a Regular Curve-Family

$Y_\alpha^-(t_0) \neq 0$ and $Z_\alpha^-(t_0) = 0$), then $d_{t_0}^\alpha$ is a *left discontinuity*. If $Y_\alpha^+(t_0) = 0$ and $Z_\alpha^+(t_0) \neq 0$ (or $Y_\alpha^-(t_0) = 0$ and $Z_\alpha^-(t_0) \neq 0$), then $d_{t_0}^\alpha$ is a *right discontinuity*. If $Y_\alpha^+(t_0) \neq 0$ and $Z_\alpha^+(t_0) \neq 0$ (or $Y_\alpha^-(t_0) \neq 0$ and $Z_\alpha^-(t_0) \neq 0$), $d_{t_0}^\alpha$ is a *bilateral discontinuity*.

Let $\varphi(t)$ be a real function of t in $0 \leqq t < \infty$. A value t_0 will be termed respectively a discontinuity, left discontinuity, right discontinuity, or bilateral discontinuity of $\varphi(t)$ according as φ is discontinuous at t_0, discontinuous at t_0

but continuous to the right at t_0, discontinuous at t_0 but continuous to the left at t_0, or discontinuous on both sides of t_0.

LEMMA. *There exist real, monotone strictly increasing functions* $\varphi_\alpha^+(t)$ *and* $\varphi_\alpha^-(t)$ *in* $0 \leqq t < \infty$ *whose discontinuities are precisely at the values t such that* d_t^α *is a discontinuity with respect to* V_α^+ *and* V_α^- *respectively and that t is a left, right, or bilateral discontinuity of* $\varphi_\alpha^+(t)$ *or* $\varphi_\alpha^-(t)$ *according as* d_t^α *is a left, right, or bilateral discontinuity with respect to* V_α^+ *and* V_α^- *respectively. Further, these functions can be so chosen that*

$$\lim_{t\to\infty} \varphi_\alpha^+(t) = \lim_{t\to\infty} \varphi_\alpha^-(t) = \infty, \qquad \varphi_\alpha^+(0) = \varphi_\alpha^-(0) = 0.$$

Proof. There are at most countably many discontinuities $d_{t_{\alpha,k}}^\alpha$ with respect to V_α^+. Number the corresponding set of values of t as t_α^i $(i = 1, 2, \cdots)$, where the range of i may be infinite, finite, or void. Then let ${}^+\varphi_\alpha^i(t) = 0$ for $0 < t < t_\alpha^i$, $= 1$ for $t_\alpha^i < t < \infty$; let ${}^+\varphi_\alpha^i(t_\alpha^i) = 1, 0$, or $\frac{1}{2}$ according as $d_{t_\alpha^i}^\alpha$ is a left, right, or bilateral discontinuity with respect to V_α^+. Now set

$$\varphi_\alpha^+(t) = t + \sum_i \frac{{}^+\varphi_\alpha^i(t)}{2^i}.$$

$\varphi_\alpha^+(t)$ then has the desired properties.

To obtain $\varphi_\alpha^-(t)$, we number the discontinuity values $t_{\alpha,k}$ with respect to V_α^- as t_α^i $(i = -1, -2, \cdots)$ and proceed in the same way. The lemma is thus established.

Suppose now that $\varphi_\alpha^+(t)$ and $\varphi_\alpha^-(t)$ are chosen as in the lemma. Then choose a sequence s_n of points on $0 \leqq t < \infty$, everywhere dense on that infinite interval, and further including all points t_α^i $(i = \pm 1, \pm 2, \cdots)$. Let ${}^+\psi_\alpha^n(t)$ $(n = 1, 2, \cdots)$ be the function equal to $\varphi_\alpha^+(t)$ for $t = 0, 1, 2, \cdots$ and for $t = s_1, s_2, \cdots, s_n$, and varying linearly between these values. Similarly, let ${}^-\psi_\alpha^n(t)$ be the function equal to $\varphi_\alpha^-(t)$ for $t = 0, 1, 2, \cdots$ and for $t = s_1, s_2, \cdots, s_n$ and varying linearly between these values. It follows from a theorem of Lebesgue[2] that $\lim_{n\to\infty} {}^+\psi_\alpha^n(t) = \varphi_\alpha^+(t)$, $\lim_{n\to\infty} {}^-\psi_\alpha^n(t) = \varphi_\alpha^-(t)$. Further, since $\varphi_\alpha^+(t)$ and $\varphi_\alpha^-(t)$ are monotone strictly increasing, ${}^+\psi_\alpha^n(t)$ and ${}^-\psi_\alpha^n(t)$ are monotone strictly increasing, and are, moreover, continuous.

Construction of curves $\tilde{D}_t^\alpha$. Corresponding to each element d_t^α we now define a curve $\tilde{D}_t^\alpha$. The curves $\tilde{D}_t^\alpha$ will fill the region $|x| < 1$, $0 \leqq y < \infty$ (which we shall later map homeomorphically on a semicircle). For each t the curve $\tilde{D}_t^\alpha$ is defined as the graph of the function $y = f_t^\alpha(x)$ in $-1 < x < 1$, where $f_t^\alpha(0) = t$, $f_t^\alpha(1 - 2^{-n}) = {}^+\psi_\alpha^n(t)$, $f_t^\alpha(-1 + 2^{-n}) = {}^-\psi_\alpha^n(t)$ $(n = 1, 2, \cdots)$ $(0 \leqq t < \infty)$ and $f_t^\alpha(x)$ varies linearly between these values. It follows immediately that $\lim_{x\to 1} f_t^\alpha(x) = \varphi_\alpha^+(t)$, $\lim_{x\to -1} f_t^\alpha(x) = \varphi_\alpha^-(t)$.

[2] See E. Borel, *Leçons sur les Fonctions de Variables Réelles*, Paris, 1928, pp. 97–98.

1.5. **Construction of curves of $V_\alpha - c_\alpha$.** The curves $\tilde{D}_t^\alpha$ have limit points $(+1, \varphi_\alpha^+(t))$ and $(-1, \varphi_\alpha^-(t))$ on $x = 1$ and $x = -1$. Since φ_α^+ and φ_α^- are monotone functions, these limit points do not in general fill the lines $x = \pm 1$, but leave gaps which are half-open intervals. If, for example, the function $\xi_\alpha = \varphi_\alpha^+(t)$ is discontinuous to the left at t_α^i, then $\lim_{t \to t_\alpha^i - 0} \varphi_\alpha^+(t) = \underline{\xi}_\alpha^i$ is less than $\xi_\alpha^i = \varphi_\alpha^+(t_\alpha^i)$. The half-open interval $[\underline{\xi}_\alpha^i, \xi_\alpha^i)$ then lies outside of the range of $\varphi_\alpha^+(t)$. But, by the above lemma and the definitions of §1.4, this occurs precisely when $Y_\alpha^+(t_\alpha^i) \neq 0$. Similarly, if $Z_\alpha^+(t_\alpha^i) \neq 0$, then there is a gap $(\xi_\alpha^i, \bar{\xi}_\alpha^i]$ in the range of $\varphi_\alpha^+(t)$, where $\bar{\xi}_\alpha^i = \lim_{t \to t_\alpha^i + 0} \varphi_\alpha^+(t)$; if $Y_\alpha^-(t_\alpha^i) \neq 0$, there is a gap $[\underline{\eta}_\alpha^i, \eta_\alpha^i)$ in the range of $\eta_\alpha = \varphi_\alpha^-(t)$; if $Z_\alpha^-(t_\alpha^i) \neq 0$, there is a gap $(\eta_\alpha^i, \bar{\eta}_\alpha^i]$ in the range of $\eta_\alpha = \varphi_\alpha^-(t)$.

If any of the sets $Y_\alpha^+(t_\alpha^i)$, $Z_\alpha^+(t_\alpha^i)$, $Y_\alpha^-(t_\alpha^i)$, $Z_\alpha^-(t_\alpha^i)$ is non-void ($i = \pm 1, \pm 2, \cdots$), we then represent its elements by corresponding curves $\tilde{C}_{\alpha,k}$ in the corresponding gap on $x = 1$ or $x = -1$ left by the range of $(1, \xi_\alpha)$ or $(-1, \eta_\alpha)$. For example, if $Y_\alpha^+(t_\alpha^i) \neq 0$, then $Y_\alpha^+(t_\alpha^i)$ is, by Theorem 3, a simply-ordered set of elements $c_{\alpha,k}$ of V_α^+. To each $c_{\alpha,k}$ we then choose an open interval $\tilde{C}_{\alpha,k}$ on $x = 1$, interior to the interval $\underline{\xi}_\alpha^i < y < \xi_\alpha^i$. Further, we carry out these choices for all $c_{\alpha,k}$ in $Y_\alpha^+(t_\alpha^i)$ in such a way that the corresponding intervals $\tilde{C}_{\alpha,k}$ are pairwise disjoint and that, if we order the $\tilde{C}_{\alpha,k}$ by the size of the y-coördinates of their midpoints, then $\tilde{C}_{\alpha,k} < \tilde{C}_{\alpha,l}$ is equivalent to $c_{\alpha,k} < c_{\alpha,l}$. This construction is possible since the $c_{\alpha,k}$ form an at most countably infinite set. It can be carried out similarly for all $Z_\alpha^+(t_\alpha^i)$, $Y_\alpha^-(t_\alpha^i)$, $Z_\alpha^-(t_\alpha^i)$ for the corresponding gaps. Since the gaps themselves do not overlap, the resulting intervals $\tilde{C}_{\alpha,k}$ will never overlap each other, and no $\tilde{C}_{\alpha,k}$ intersects a $\tilde{D}_t^\alpha$. We shall further assume the $\tilde{C}_{\alpha,k}$ chosen less in length than constants $\epsilon_{\alpha,k} > 0$. The precise value of the $\epsilon_{\alpha,k}$ will be indicated below.

1.6. **Further information on the curves $\tilde{D}_t^\alpha$ and $\tilde{C}_{\alpha,k}$.**

THEOREM 5. *For each fixed α the curves $\tilde{D}_t^\alpha$ fill the region $-1 < x < 1$, $0 \leqq y < \infty$.*

Proof. Consider the half-strip $0 \leqq x \leqq \frac{1}{2}$, $0 \leqq y < \infty$. The curves $\tilde{D}_t^\alpha$ therein join the points $(0, t)$ and $(\frac{1}{2}, {}^+\psi_\alpha^1(t))$ by straight lines. ${}^+\psi_\alpha^1(t)$ is monotone increasing and $\lim_{t \to \infty} {}^+\psi_\alpha^1(t) = \infty$, ${}^+\psi_\alpha^1(0) = 0$. The equation of each line can be written as

$$t(1 - 2x) + 2\,{}^+\psi_\alpha^1(t)x = y. \tag{1}$$

But for fixed x in $0 \leqq x \leqq \frac{1}{2}$, the left side of this equation is a monotone increasing continuous function of t which $\to \infty$ as $t \to \infty$ and $= 0$ for $t = 0$. Hence it takes on each value y in $0 \leqq y < \infty$ just once. This shows that through each point of this half-strip passes one and only one curve. The same argument applies to each half-strip $1 - 2^{-n} \leqq x \leqq 1 - 2^{-n-1}$, $0 \leqq y < \infty$, where each line on a $\tilde{D}_t^\alpha$ joins $(0, \tau_n(t))$, where $\tau_n = {}^+\psi_\alpha^n(t)$, to $(0, \tau_{n+1}(t))$, where $\tau_{n+1} = {}^+\psi_\alpha^{n+1}(t)$, since τ_{n+1} is a monotone increasing function of τ_n, and similarly to

each half-strip $-1 + 2^{-n-1} \leqq x \leqq -1 + 2^{-n}$. Hence through each point of the region $-1 < x < 1, 0 \leqq y < \infty$ passes one and only one $\tilde{D}_t^a$. This completes the proof.

Lemma. *Let a curve-family Φ be given by curves $C_u : \omega = g(\zeta, u)$ $(a \leqq \zeta \leqq b;$ $c \leqq u \leqq d)$ in the (ζ, ω)-plane, where $g(\zeta, u)$ is for each u continuous in ζ, and $g(a, u) = u$. Let Φ fill the closed region Γ: $a \leqq \zeta \leqq b$, $g(\zeta, c) \leqq \omega \leqq g(\zeta, d)$. Then there is a homeomorphism defined on Γ, leaving ζ invariant, taking the C_u onto parallel lines ω = constant.*

Proof. The equation $\omega = g(\zeta, u)$ determines u as a single-valued continuous function of ζ and ω in Γ. For through each point (ζ_0, ω_0) passes a unique C_u. Further, given any $\epsilon > 0$, the curves C_u with $| u - u_0 | < \epsilon$ fill a strip containing (ζ_0, ω_0) in its interior. Hence, for δ sufficiently small and > 0, $| \zeta - \zeta_0 | < \delta$, $| \omega - \omega_0 | < \delta$ implies $| u - u_0 | < \epsilon$. Thus u is a single-valued continuous function $h(\zeta, \omega)$.

We then make the transformation on Γ: $\zeta' = \zeta$, $\omega' = h(\zeta, \omega)$. This transformation is one-to-one and continuous, hence a homeomorphism, and takes Γ onto Γ': $a \leqq \zeta' \leqq b, c \leqq \omega' \leqq d$; each curve C_u becomes the curve $C'_u : \omega' = u$. The lemma thus follows.

Theorem 6. *To each closed interval I lying on $\tilde{D}_0^a$ or on a $\tilde{C}_{\alpha,k}$ corresponds a point set G lying in $-1 < x < 1$, $0 < y < \infty$ and such that the set $G \cup I = H$ can be mapped homeomorphically on a region $a \leqq \zeta \leqq b$, $c \leqq \omega \leqq d$ so that the inverse image of each line ζ = constant, $\neq a$, is an arc of a curve $\tilde{D}_t^a$, of $\zeta = a$ is I. Further, if I is on $\tilde{D}_0^a$, and (x_0, y_0) is any point such that $(x_0, 0)$ is interior to I and $y_0 > 0$, then G can be chosen to include (x_0, y_0) as interior point.*

Proof. If I lies on $\tilde{D}_0^a$, let I be given by $a \leqq x \leqq b$, $y = 0$. Let (x_0, y_0) be given with $a < x_0 < b$ and $y_0 > 0$. Let $\tilde{D}_{t_0}^a$ be the curve through (x_0, y_0) in virtue of Theorem 5. Take $t_1 > t_0$ and let G be the region $a \leqq x \leqq b, 0 < y \leqq f_{t_1}^a(x)$. Through each point of $H = G \cup I$ then passes one and only one $\tilde{D}_t^a$. The conditions of the above lemma hold and the desired homeomorphism is obtained.

If I lies on a $\tilde{C}_{\alpha,k}$, then suppose, for example, that $\tilde{C}_{\alpha,k}$ is in the interval $\underline{\xi}_\alpha^i < y < \xi_\alpha^i$ on $x = 1$ and that I is the subinterval $y_0 \leqq y \leqq y_1$. For N sufficiently large we then have $s_N = t_\alpha^i$ and ${}^+\psi_\alpha^N(t_\alpha^i) = \xi_\alpha^i$. Let m be the largest integer less than t_α^i and, if any of the numbers $s_1, s_2, \cdots, s_{N-1}$ fall in the interval (m, t_α^i), let s_{n_1} denote the largest one. Otherwise let $s_{n_1} = m$. Let for $j = 2, 3, \cdots$ s_{n_j} be the s_n of smallest index $n_j > N$ to fall in the interval $(s_{n_{j-1}}, t_\alpha^i)$. We have then $s_{n_j} < s_{n_{j+1}}$ and, since the s_n are everywhere dense, $\lim_{j \to \infty} s_{n_j} = s_N = t_\alpha^i$. Further, for $n > N$ we set

$$h_n = \max\, (\max\, ({}^+\psi_\alpha^{n+1}(t) - {}^+\psi_\alpha^n(t)), 0)$$

for t restricted so that $\underline{\xi}_\alpha^i \leqq {}^+\psi_\alpha^{n+1}(t) \leqq \xi_\alpha^i$. For $n + 1$ not equal to an n_j we must have $h_n = 0$, since ${}^+\psi_\alpha^{n+1}(t) = {}^+\psi_\alpha^n(t)$ in the interval $s_{n_{j_0}} \leqq t \leqq t_\alpha^i$, where

n_{j_0} is the largest $n_j < n + 1$. This follows from the very definition of ${}^+\psi_\alpha^n(t)$ and the fact that

$${}^+\psi_\alpha^n(s_{n_{j_0}}) = \varphi_\alpha^+(s_{n_{j_0}}) < \xi_\alpha^i.$$

If $n + 1 = n_{j+1}$, we have

$$0 \leqq h_n \leqq {}^+\psi_\alpha^{n+1}(s_{n_{j+1}}) - {}^+\psi_\alpha^n(s_{n_j}) = \varphi_\alpha^+(s_{n_{j+1}}) - \varphi_\alpha^+(s_{n_j}).$$

It follows that

$$\begin{aligned}\sum_{N+1}^{\infty} h_n &\leqq \sum_{j=1}^{\infty} [\varphi_\alpha^+(s_{n_{j+1}}) - \varphi_\alpha^+(s_{n_j})] \\ &\leqq \lim_{j\to\infty} [\varphi_\alpha^+(s_{n_{j+1}}) - \varphi_\alpha^+(s_{n_1})] = \xi_\alpha^i - \varphi_\alpha^+(s_{n_1}).\end{aligned} \tag{1}$$

Finally, if we refer to the curves $\tilde{D}_t^\alpha$, we see that $h^n/2^{n+1}$ is an upper bound for the slope of the curves $\tilde{D}_t^\alpha$ in the rectangle

$$1 - 2^{-n} \leqq x \leqq 1 - 2^{-n-1}, \underline{\xi}_\alpha^i \leqq y \leqq \xi_\alpha^i.$$

Now let $g(x)$ be the function equal to $\sum_{n=N+1}^{r} (h_n + 2^{-n})$ for $x = 1 - 2^{-r-1}$ and varying linearly between these values. If we take $g(1)$ as $\lim_{x\to 1-0} g(x)$ (which exists, by (1)), then $g(x)$ becomes monotone increasing and continuous in the closed interval $1 - 2^{-N-1} \leqq x \leqq 1$.

Next, for any point $(1, \xi)$ of the interval I let $g_\xi(x) = g(x) - g(1) + \xi$. Hence $\lim_{x\to 1-0} g_\xi(x) = \xi$. Then for n sufficiently large $\underline{\xi}_\alpha^i < g_\xi(x) < \xi_\alpha^i$ for $1 - 2^{-n} \leqq x \leqq 1$ and $(1, \xi)$ on I. The curve $y = g_\xi(x)$ then meets each curve $\tilde{D}_t^\alpha$ at most once. For the slope of $y = g_\xi(x)$ is always greater than that of a $\tilde{D}_t^\alpha$ at any point of the rectangle.

If we now apply the above lemma to the curves $y = g_\xi(x)$ in the region Γ which they fill for $(1, \xi)$ in I and $1 - 2^{-n} \leqq x \leqq 1$, then a homeomorphism T_1 leaving x invariant maps them on the lines $y =$ constant, for y in I. Under T_1 each part of a curve $\tilde{D}_t^\alpha$ in Γ becomes a curve meeting each line $y =$ constant at most once, hence a curve of the form $x = \theta(y)$. In particular, the curve $x = \theta^*(y)$ through $x = 1 - 2^{-n}$, $y = y_1$ is defined for all y in I. Through each point of the region $\Gamma_1 : \theta^*(y) \leqq x \leqq 1$, $y_0 \leqq y \leqq y_1$ then passes a curve $x = \theta(y)$ or the curve $x = 1$ (i.e., the interval I). A second application of the lemma therefore gives a homeomorphism T_2 under which the curves $x = \theta(y)$ and I become the lines $x =$ constant. Set $\zeta = -x$, $\omega = y$, and the theorem is established.

1.7. **Construction of the family** F. Thus far we have indicated the construction of a set of curves for each $\lambda(V_\alpha)$ of E. This set consists of a set of curves $\tilde{D}_t^\alpha$ filling the half-strip $-1 < x < 1$, $0 \leqq y < \infty$ and a set of intervals $\tilde{C}_{\alpha,k}$ on the boundaries $x = 1$ and $x = -1$. We shall now map each such half-strip

on a certain region of a circle. By doing this successively for $\lambda(V_1)$, all $\lambda(V_{1,k})$, all $\lambda(V_{1,k,l})$, $\cdots$, we shall obtain a family of curves filling a semicircle.

For $\lambda(V_1)$ we perform the mapping in two stages, as follows. We first perform the homeomorphism $T_1 : x' = x(1 + y^2)^{-\frac{1}{2}}$, $y' = y(1 + y^2)^{-\frac{1}{2}}$ on the half-strip $-1 \leqq x \leqq 1$, $0 \leqq y < \infty$. This maps the half-strip on the semicircle $0 \leqq y' \leqq (1 - x'^2)^{\frac{1}{2}}$ minus the point $x' = 0$, $y' = 1$. The line $-1 \leqq x \leqq 1$, $y = 0$ becomes the diameter $-1 \leqq x' \leqq 1$, $y' = 0$. The lines $x = +1$ and $x = -1$ become the arcs $0 \leqq \theta' < \frac{1}{2}\pi$, $r' = 1$ and $\frac{1}{2}\pi < \theta' \leqq \pi$, $r' = 1$ in the corresponding polar coördinates. Hence the curves $\bar{D}_t^1$ have as images curves $\bar{D}_t^1$ joining points of these two arcs. The curves $\bar{C}_{1,k}$ have as images open subarcs $\bar{C}_{1,k}$ of the two arcs. Let $C_{1,k}$ denote the corresponding chords, minus end-points.

We now perform a further homeomorphism. The chords $C_{1,k}$, plus the points of the semicircle not on the $\bar{C}_{1,k}$ form an arc $r' = g_1(\theta')$ joining the two ends of the diameter. Our homeomorphism is now taken as $T_2 : \theta = \theta'$, $r = r'g_1(\theta')$. This maps the semicircle plus interior on the set $0 \leqq r \leqq g_1(\theta)$, $0 \leqq \theta \leqq \pi$. The arcs $\bar{C}_{1,k}$ become the chords $C_{1,k}$, the diameter $\bar{C}_1$ (or $\bar{D}_0^1$) remains fixed, becomes the curve C_1, the curves $\bar{D}_t^1$ become curves D_t^1 joining points of the circle $r = 1$.

We now represent each $\lambda(V_{1,k})$ in the same way as a family of curves in the segment bounded by $C_{1,k}$, the elements $c_{1,k,l}$ becoming chords $C_{1,k,l}$ in the segment. $\lambda(V_{1,k})$ is thus represented by curves filling a region $g_1(\theta) \leqq r \leqq g_{1,k}(\theta)$, where θ is restricted to the interval determined by $C_{1,k}$ and $r < 1$.

Proceeding in this way, we fit curves corresponding to each $\lambda(V_\alpha)$ in the semicircle, the c_α being represented by chords C_α. A similar process is carried out for the $\lambda(V_\alpha^*)$ in the lower semicircle. In all cases the homeomorphism of the half-strip $-1 \leqq x \leqq 1$, $0 \leqq y < \infty$ is analogous to that for $\lambda(V_1)$ above, except that in all cases it must be chosen to preserve orientation.

We thus obtain a family F of curves C in one-to-one correspondence with the elements c of E (see Theorem I.39). It remains to be shown that F actually fills the interior of the circle, that F is regular, and that F is isomorphic to E.

1.8. Proof that F is a regular family filling the interior of the circle. We first remark that the numbers $\epsilon_{\alpha,k}$ and $\epsilon_{\alpha,k}^*$ (see §1.5 above) have not yet been fixed.

THEOREM 7. *For proper choice of the $\epsilon_{\alpha,k}$ and $\epsilon_{\alpha,k}^*$, F fills the interior of the circle.*

Proof. Denote by $G_{\alpha,k}$ the region $g_\alpha(\theta) \leqq r \leqq g_{\alpha,k}(\theta)$ ($r < 1$, θ on the arc determined by $C_{\alpha,k}$) in which are the curves of F corresponding to $\lambda(V_{\alpha,k})$ and by G_1 the region $0 \leqq r \leqq g_1(\theta)$ in which lie the curves corresponding to $\lambda(V_1)$. Similarly denote by $G_{\alpha,k}^*$ the region $g_\alpha^*(\theta) \leqq r \leqq g_{\alpha,k}^*(\theta)$ corresponding to $\lambda(V_\alpha^*)$ and by G_1^* the region $0 \leqq r \leqq g_1^*(\theta)$ corresponding to $\lambda(V_1^*)$. The curves of F actually fill the regions G_α and G_α^*, as follows from Theorem 5.

We now choose the $\epsilon_{\alpha,k}$ and $\epsilon_{\alpha,k}^*$ so small that the chords $C_{\alpha,k}$ and $C_{\alpha,k}^*$ are so

small that $g_\alpha(\theta) > 1 - 2^{-m}$, $g_\alpha^*(\theta) > 1 - 2^{-n}$, where m, n respectively are the numbers of elements in the two sequences α. We must now show that the set $\sum G_\alpha \cup \sum G_\alpha^*$ coincides with the interior of the circle.

Suppose the G_α did not fill the interior of the upper semicircle. Let P be a point of that interior and not in $\sum G_\alpha$. Since P is not in G_1, it lies in some segment bounded by a chord $C_{1,k}$. Repeating this reasoning we find that there is a sequence k_n $(n = 1, 2, \cdots)$ such that P is in the smaller segment bounded by $C_{1,k_1,k_2,\cdots,k_n}$ for every n. Since $g_\alpha(\theta) > 1 - 2^{-n}$, this implies that the distance of P from the origin is greater than $1 - 2^{-n}$ for every n. This is impossible if P is interior to the semicircle. Hence the set $\sum G_\alpha$ covers the interior of the upper semicircle, and similarly the set $\sum G_\alpha^*$ covers the interior of the lower semicircle.

THEOREM 8. *F is regular.*

Proof. Consider an interior point P of a G_α. From Theorem 6 we conclude that there is an r-neighborhood of P (one of whose sides is an interval I on C_α).

If P is on a $C_{\alpha,k}$, then choose I, an interval of $C_{\alpha,k}$, to contain P. There is then, by Theorem 6, an r-neighborhood U_1 in G_α one of whose sides is I and one U_2 in $G_{\alpha,k}$, one of whose sides is I. From the construction of F, $\bar{U}_1$ and $\bar{U}_2$ have only I in common. Hence $\bar{U}_1 \cup \bar{U}_2$ is an r-neighborhood.

The cases when P is on C_1 or on the lower semicircle are handled similarly. Hence F is regular.

1.9. **The chordal relations in F.** The family F has been constructed as a set of curves filling the interior of a circle. Each curve C has exactly two limit points, which are distinct and lie on the circumference. No two curves have a common limit point. If we regard the interior of the circle as homeomorphic image of the plane, then we can introduce the chordal relations in F just as in the case of a family filling the plane. Let T be a fixed o-homeomorphism of the interior of the circle on the plane. For each triple C_1, C_2, C_3 of F we then assign the chordal relations of $T(C_1)$, $T(C_2)$, $T(C_3)$.

THEOREM 9. *If C_i is a curve of F with limit points P_i and Q_i $(i = 1, 2, 3)$, then $C_2 \mid C_1 \mid C_3$ is equivalent to the condition that the limit points, if properly named, lie in the order P_3, P_1, P_2, Q_2, Q_1, Q_3 on the circumference. $\mid C_1, C_2, C_3 \mid^+$ is equivalent to the condition that the limit points can be so named that they lie in the order P_1, Q_1, P_2, Q_2, P_3, Q_3 on the circle, and that this order determines a positive orientation of the circle.*

Proof. Suppose $C_2 \mid C_1 \mid C_3$. Then $T(C_2) \subset \mathfrak{D}(T(C_1))$, $T(C_3) \subset \mathfrak{D}^*(T(C_1))$ for proper choices. Let $\Delta(C_1)$, $\Delta^*(C_1)$ be the sets $T^{-1}(\mathfrak{D}(T(C_1)))$, $T^{-1}(\mathfrak{D}^*(T(C_1)))$. Then $C_2 \subset \Delta(C_1)$, $C_3 \subset \Delta^*(C_1)$. If we name P_1 and Q_1 in a fixed way, then C_2 must lie on one arc P_1Q_1, C_3 on the other. Hence the other limit points can be so named that the final order is P_1, P_2, Q_2, Q_1, Q_3, P_3, as desired. Conversely,

if such a naming is possible, then $C_2 \subset \Delta(C_1)$, $C_3 \subset \Delta^*(C_1)$, whence $T(C_2) \mid T(C_1) \mid T(C_3)$ and $C_2 \mid C_1 \mid C_3$.

Suppose $| C_1, C_2, C_3 |^+$. There is then a positively oriented closed curve $T(M_1)T(M_2)T(M_3)T(M_1)$ through the points $T(M_1)$ on $T(C_1)$, $T(M_2)$ on $T(C_2)$, $T(M_3)$ on $T(C_3)$ and not meeting $T(C_1)$, $T(C_2)$, $T(C_3)$ otherwise. (See I, §2.5.) The inverse image of this curve is a positively oriented curve $M_1M_2M_3M_1$ meeting C_1, C_2, C_3 only at M_1, M_2, M_3. The existence of such a curve implies that C_1, C_2, C_3 determine circular arcs P_1Q_1, P_2Q_2, P_3Q_3 which do not overlap and which, in the order given, follow the positive orientation of the circle. The limit points can thus be named as stated in the theorem. Conversely, if they can be so named, then $C_1 \mid C_2 \mid C_3$, $C_3 \mid C_1 \mid C_2$, $C_2 \mid C_1 \mid C_3$ are all impossible by the first part of the theorem. Hence $| C_1, C_2, C_3 |^{\pm}$. But $| C_1, C_2, C_3 |^-$ implies by the previous reasoning that C_1, C_2, C_3 determine circular arcs P_1Q_1, P_2Q_2, P_3Q_3 following the negative orientation of the circle. This contradicts the given naming of the limit points. Hence $| C_1, C_2, C_3 |^+$ and the theorem is established.

THEOREM 10. *The sets* $W_\alpha = C_\alpha \cup \sum C_{\alpha,k}$ *and* $W_\alpha^* = C_\alpha^* \cup \sum C_{\alpha,k}^*$ *determine a normal subdivision of* F.

Proof. Consider first the curves of the upper semicircle. These form a set $F_0 = C_1 \cup \delta(C_1)$, by the definition of $\delta(C_1)$ (I, §3.2) and the preceding theorem. The curves C_α and $C_{\alpha,k}$ of W_α are chords of the circle and are the images of the line $y = 0$ and of intervals on $x = 1$, $x = -1$ of the half-strip $-1 \leqq x \leqq 1$, $0 \leqq y < \infty$. It follows that no one of these chords separates any two others. Hence, by the preceding theorem, $| C_1, C_2, C_3 |^{\pm}$ for any three curves of W_α. Thus each W_α is cyclic.

For each $C_{\alpha,k}$, $\delta(C_{\alpha,k})$ denotes the curves of one of the two segments bounded by $C_{\alpha,k}$. We shall fix the choice by requiring that $\delta(C_{\alpha,k})$ does not include C_α. We then have

$$\theta(W_\alpha) = \delta(C_\alpha) \cdot \prod \delta^*(C_{\alpha,k}),$$

where this equation serves as definition of $\theta(W_\alpha)$ in case W_α contains only C_α. It immediately follows that $\lambda(W_\alpha)$ and $\lambda(W_{\alpha,k})$ are adjacent.

Each set $\lambda(W_\alpha)$ is homeomorphic image of the curves $\bar{D}_t^\alpha$ and $\bar{C}_{\alpha,k}$ in $-1 \leqq x \leqq 1$, $0 \leqq y < \infty$. It follows from Theorem I.29 that $D_{t_1}^\alpha \mid D_{t_2}^\alpha \mid D_{t_3}^\alpha$ is equivalent to $t_1 < t_2 < t_3$ or $t_3 < t_2 < t_1$. Hence the D_t^α in $0 \leqq t < \infty$ are half-parallel.

Since each $\theta(W_\alpha)$ fills the interior of G_α, we must have $F_0 = \sum \lambda(W_\alpha)$.

Conditions (1)-(5) of seminormality are thus verified. To verify condition (6) we take any curve $C_{\alpha,k}$ of $W_\alpha - C_\alpha$. By our construction of the curves $\bar{D}_t^\alpha$ and $\bar{C}_{\alpha,k}$, the interval $\bar{C}_{\alpha,k}$ lies below the limit point of $\bar{D}_{t_{\alpha,k+1}}^\alpha$ on $x = 1$ or $x = -1$ (according to the case). Hence no one of the three curves C_α, $C_{\alpha,k}$, $D_{t_{\alpha,k+1}}^\alpha$ can separate the other two. Hence, by Theorem 9, $| C_\alpha, C_{\alpha,k}, D_{t_{\alpha,k+1}}^\alpha |^{\pm}$.

Condition (6) thus holds. Hence F_0 is seminormal and similarly $F_0^* = C_1 \cup \delta^*(C_1)$ is seminormal. Thus F is normal.

THEOREM 11. *E is isomorphic to F.*

Proof. The elements C of F have been constructed to be in one-to-one correspondence with those of E. It remains to verify that this correspondence preserves the chordal relations. By Theorem 1, we need only show that the relations are preserved in each $\lambda(V_\alpha)$. By Theorem 4, we need only show that in each $\lambda(V_\alpha)$ and corresponding $\lambda(W_\alpha)$ the relations are the same for corresponding triples of types (α) or (β).

Consider a triple of type (α). Suppose, for example, $c_{\alpha,k} \subset V_\alpha^+$. The interval $\tilde{C}_{\alpha,k}$ then lies on $x = 1$. Further, a limit point of a curve $\tilde{D}_t^\alpha$ on $x = 1$ lies above $\tilde{C}_{\alpha,k}$ according as $d_t^\alpha \subset B(c_{\alpha,k})$, i.e., $| c_\alpha , c_{\alpha,k} , d_t^\alpha |^+$. But the fact that $\tilde{D}_t^\alpha$ has limit point on $x = 1$ above $\tilde{C}_{\alpha,k}$ implies that the limit points of the curves C_α, $C_{\alpha,k}$, D_t^α, in that order, follow the positive orientation on the circle. Hence, by Theorem 9, $| C_\alpha , C_{\alpha,k} , D_t^\alpha |^+$. If, on the other hand, the limit point of $\tilde{D}_t^\alpha$ on $x = 1$ lies below $\tilde{C}_{\alpha,k}$, then $d_t^\alpha \not\subset B(c_{\alpha,k})$, and hence $c_\alpha \mid d_t^\alpha \mid c_{\alpha,k}$. The image D_t^α of $\tilde{D}_t^\alpha$ on the circle will then separate C_α and $C_{\alpha,k}$, whence $C_\alpha \mid D_t^\alpha \mid C_{\alpha,k}$. Thus for triples of type (α) the relations are the same.

For a triple of type (β): $c_{\alpha,k}$, $c_{\alpha,l}$, $c_{\alpha,m}$ assume, for example, that $k > 0$, $l > 0$, $m > 0$ and that all are in the same set $Y_\alpha^+(t_0)$. They are then represented on $x = 1$ by intervals $\tilde{C}_{\alpha,k}$, $\tilde{C}_{\alpha,l}$, $\tilde{C}_{\alpha,m}$ in the same order as the $c_{\alpha,k}$, $c_{\alpha,l}$, $c_{\alpha,m}$; that is, if $c_{\alpha,k} < c_{\alpha,l} < c_{\alpha,m}$, then $\tilde{C}_{\alpha,k}$ lies below $\tilde{C}_{\alpha,l}$, which lies below $\tilde{C}_{\alpha,m}$. By Theorem 3, $| c_{\alpha,k} , c_{\alpha,l} , c_{\alpha,m} |^+$. By the construction of F, the chords $C_{\alpha,k}$, $C_{\alpha,l}$, $C_{\alpha,m}$ follow the positive orientation of the circle, whence $| C_{\alpha,k} , C_{\alpha,l} , C_{\alpha,m} |^+$. Thus, in this case, the relations of type (β) are invariant, and the same reasoning holds for the other cases.

We thus conclude that each $\lambda(V_\alpha)$ is isomorphic to the corresponding $\lambda(W_\alpha)$ and similarly that each $\lambda(V_\alpha^*)$ is isomorphic to the corresponding $\lambda(W_\alpha^*)$. Hence E is isomorphic to F.

We now finally obtain

THEOREM 12. *Corresponding to every normal chordal system E there exists a regular curve-family F filling the plane with CS(F) isomorphic to E.*

The family F here is the image of the above family F under the homeomorphism T of the interior of the circle on the plane.

2. O-equivalence of isomorphic curve-families

2.1. **The equivalence and o-equivalence classes of curve-families.** Let F_1 and F_2 be two regular curve-families filling the plane. F_1 is termed *equivalent* to F_2 if there is a homeomorphism of the plane onto itself transforming each curve of F_1 onto a curve of F_2. This equivalence is reflexive, symmetric and transitive. Hence the curve-families are grouped in *equivalence classes*.

We now make a finer subdivision. F_1 is termed *o-equivalent* to F_2 if there is an *o-homeomorphism* of the plane onto itself transforming each curve of F_1

onto a curve of F_2. The o-equivalence also divides the set of families into classes, the *o-equivalence classes.* Each equivalence class is then the union of two (possibly coinciding) o-equivalence classes.

THEOREM 13. *If F_1 is o-equivalent to F_2, then $CS(F_1)$ is isomorphic to $CS(F_2)$.*

For the chordal relations, as defined in I, §§2.4 and 2.5, are invariant under an o-homeomorphism.

In the present section we shall establish the converse of Theorem 13: if $\mathrm{CS}(F_1)$ is isomorphic to $\mathrm{CS}(F_2)$, then F_1 is o-equivalent to F_2. From Theorem I.38 and Theorem 12 it will then follow that the o-equivalence classes are in one-to-one correspondence with the isomorphism classes of normal chordal systems.

2.2. Map of $\theta(V)$ onto parallel lines. We suppose given two regular curve-families F_1 and F_2 filling the plane, and further that $E^1 = \mathrm{CS}(F_1)$ is isomorphic to $E^2 = \mathrm{CS}(F_2)$. Let $f(C^1) = C^2$ be a fixed isomorphism of E^1 onto E^2.

By Theorem I.38, E^1 is normal. Suppose then E^1 divided into two subsets $E_0^1 = C_1^1 \cup \mathfrak{D}(C_1^1)$ and $E_0^{1*} = C_1^1 \cup \mathfrak{D}^*(C_1^1)$ and that E_0^1 is seminormally subdivided by cyclic subsets V_α^1 for α in A and E_0^{1*} by cyclic subsets V_α^{1*} for α in A^*.

It then immediately follows that $f(E_0^1) = E_0^2 = C_1^2 \cup \mathfrak{D}(C_1^2)$ is seminormally subdivided by the subsets $f(V_\alpha^1) = V_\alpha^2$, since seminormality is defined wholly in terms of the chordal relations, and f preserves these relations. Similarly $f(E_0^{1*}) = E_0^{2*} = C_1^2 \cup \mathfrak{D}^*(C_1^2)$ is seminormally subdivided by the sets $V_\alpha^{2*} = f(V_\alpha^{1*})$.

For simplicity, we shall from this point on use only those indices which are necessary to distinguish the elements or sets involved. Thus V will stand for any fixed one of the sets V_α^1, V_α^{1*}, V_α^2, V_α^{2*}. A pair V^1, V^2 will mean a pair in which all other indices coincide.

Consider then a fixed $\lambda(V_\alpha) = C_\alpha \cup \sum_k C_{\alpha,k} \cup \sum_t D_t^\alpha$. Omitting the index α, we write this as $\lambda(V) = C_0 \cup \sum_k C_k \cup \sum_t D_t$ (with $D_0 = C_0$). We assume the C_k numbered as in Part 1, so that $k > 0$ implies $C_k \subset V^+$, $k < 0$ implies $C_k \subset V^-$.

THEOREM 14. *There is an extended cross-section Γ from a point P_0 of C_0 to ∞ in $\theta(V)$, crossing all D_t of $\lambda(V)$ and no other curves of F.*

Proof. If V has more than one element, then choose $\mathfrak{D}(C_0)$ to include $\theta(V)$ and each $\mathfrak{D}(C_k)$ to include $\theta(V)$. Then $\theta(V) = \mathfrak{D}(C_0)\cdot\prod \mathfrak{D}(C_k)$. Otherwise $\theta(V) = \mathfrak{D}(C_0)$ for one choice of $\mathfrak{D}(C_0)$. Any curve C of $\mathfrak{D}(C_0)$ which can be joined to the point $P_0 \subset C_0$ by a curve not meeting $V - C_0$ is in $\theta(V)$. This is trivial if $V - C_0$ is void. If there is a C_k in $V - C_0$, then neither $C \mid C_0 \mid C_k$ nor $C_0 \mid C_k \mid C$ is true. Hence, by Theorem I.28, $C \subset \theta(V)$.

By Theorem I.32 the curves C_k, if infinite in number, tend uniformly to infinity. Hence a small cross-section from P_0 to a point Q_1 of $\mathfrak{D}(C_1)$ will meet no C_k and thus Q_1 lies on a curve D_{t_1} of $\theta(V)$, $t_1 > 0$. Let then τ be the least

upper bound of all values t such that there is a cross-section P_0Q with Q on D_t.

Suppose $\tau < \infty$. Then take Q_τ on C_τ. A sufficiently small neighborhood of Q_τ will be in $\mathfrak{D}(C_0)$ and in all $\mathfrak{D}(C_k)$, hence in $\theta(V)$. Thus we can draw a cross-section $Q_2Q_\tau Q_3$ in $\theta(V)$, with Q_2 on D_{t_2}, Q_2 on D_{t_3}, and, for proper numbering, $t_2 < \tau < t_3$. (See Theorem I.29.) Further, if $t_2 < t < \tau$, then $D_{t_2} \mid D_t \mid D_\tau$, whence D_t crosses $Q_2Q_\tau Q_3$. By definition of τ, there is a t_4 with $t_2 < t_4 < \tau$ such that there is a cross-section P_0Q_4, with Q_4 on D_{t_4}. There is then, by the corollary to Theorem I.30, a cross-section Q_4Q_τ. Since $C_0 \mid D_{t_4} \mid D_{t_3}$, $P_0Q_4Q_\tau Q_3$ is a cross-section. This contradicts the assumption that τ is the least upper bound. Hence $\tau = \infty$.

We can thus choose a sequence $0 < t_1 < t_2 < t_3 < \cdots < t_n < \cdots$ with $\lim_{n\to\infty} t_n = \infty$ such that there is a cross-section P_0Q_n with Q_n on D_{t_n}. By the corollary to Theorem I.30, there are then cross-sections P_0Q_1, Q_1Q_2, $\cdots$, Q_nQ_{n+1}, $\cdots$. These together form an extended cross-section $\Gamma = P_0Q_1Q_2 \cdots Q_nQ_{n+1} \cdots$ in $\theta(V) \cup C_0$. By Theorem I.29 Γ must cross all D_t in $0 \leqq t < \infty$, hence meets only those curves.

Γ must tend to infinity. For if Γ has a limit point, it cannot be on a D_t, must hence be on a C_k. By condition (6) of normality, $\mid C_0, C_k, D_t \mid^{\pm}$ for t greater than some fixed t'. But $C_0 \mid D_t \mid D_{t+s}$ for $s > 0$, whence $C_k \mid D_t \mid D_{t+s}$, and hence Γ could have no limit point on C_k. Hence Γ tends to infinity.

COROLLARY 1. *$\theta(V)$ forms an open simply-connected point set whose boundary is V.*

Proof. By the above theorem and Theorem I.37, $\theta(V)$ is an open simply-connected set whose boundary is cyclic. Since $\theta(V) = \mathfrak{D}(C_0)\cdot\prod \mathfrak{D}(C_k)$, the boundary of $\theta(V)$ must consist precisely of the curves of V.

COROLLARY 2. *Theorem* I.41 *holds for the* $\lambda(V)$ *under consideration here.*

For the proof, as given in I, depends only on the facts of Corollary 1 and on the fact that the set $\lambda(V)$ is obtained from a normal subdivision.

Remark. It can be further shown that Γ tends properly to infinity and hence, from Corollary 1, that any normal subdivision of a curve-family is obtained by the method of I.

THEOREM 15. *There exists an o-homeomorphism T mapping $C_0 \cup \theta(V)$ on the half-strip $-1 < x < 1$, $0 \leqq y < \infty$ so that each curve D_t is transformed onto the line $y = t$.*

Proof. Choose Γ as in Theorem 14. Γ is a curve $x = x(t)$, $y = y(t)$ in $0 \leqq t < \infty$. We can extend Γ to a larger extended cross-section Γ', on which the parameter t runs from -1 to ∞. The theorem is then proved in the same way as Theorem I.30. We first obtain a map onto the parallel lines filling the half-plane $y \geqq 0$, and an elementary transformation reduces this to the half-strip desired.

We remark that it is possible by this theorem to map each $\theta(V^1)$ on the cor-

responding $\theta(V^2)$ o-homeomorphically so that the curves D_t^1 become the curves D_t^2. However, this homeomorphism will not in general transform the boundary curves C_k^1 onto the corresponding boundary curves C_k^2. In the following paragraphs we shall ensure that the homeomorphism can be so chosen as to be extensible to the boundary.

2.3. **Isolation of the curves C_k.** We denote here by t_k the value of t previously denoted by $t_{a,k}$.

THEOREM 16. *A correspondence can be set up between the curves C_k and curves H_k, with the following properties:* (a) *to each C_k of V corresponds one H_k*; (b) *each H_k is an open curve, tending to infinity in both directions, and lying wholly in $\theta(V)$*; (c) $H_k \cdot H_{k'} = 0$; (d) $H_k \cdot D_{t_k} = 0$; (e) $H_k \cdot D_{t_k+1} = 0$; (f) $C_k \mid H_k \mid C_k$; (g) $C_k \mid H_k \mid H_{k'}$; (h) $C_k \mid H_k \mid D_{t_k}$; (i) $C_k \mid H_k \mid D_{t_k+1}$; (j) $C_k \mid H_k \mid C_0$; (k) $[C_0, H_k, D_{t_k+1}] \sim [C_0, C_k, D_{t_k+1}]$; (l) $[H_k, H_{k'}, H_{k''}] \sim [C_k, C_{k'}, C_{k''}]$; (m) $[C_0, H_k, H_{k'}] \sim [C_0, C_k, C_{k'}]$; (n) $[D_{t_k+1}, H_k, H_{k'}] \sim [C_0, C_k, C_{k'}]$. *Throughout k, k′, and k″ are assumed distinct.*

(*Remark.* The chordal relations have meaning in (f) $\cdots$ (n) by virtue of (b) $\cdots$ (e) and Theorem I.25 as applied to general families of non-intersecting curves.)

Proof. For convenience we renumber the C_k temporarily as a simple sequence $C_{p'}$ ($p' = 1, 2, \cdots$). We define H_1 thus: By applying a suitable o-homeomorphism of the plane onto itself we can assume that C_1 is the line $y = 0$, $-\infty < x < \infty$, and that $\theta(V)$ lies in $\mathfrak{D}(C_1)$ taken as $y > 0$. The curves C_p tend uniformly to infinity. It follows that the distance of the point $(x, 0)$ on C_1 from the set

$$(V - C_1) \cup D_{t_1} \cup D_{t_1+1}$$

is positive, equal to a number $r(x) > 0$, where $r(x)$ is a continuous function of x. H_1 is then chosen as the curve $y = \frac{1}{2}r(x)$. The curve H_1 in the original plane is then obtained by applying the inverse of the homeomorphism.

Suppose we have defined the curves H_p for $p = 1, 2, \cdots, n$. Then, as above, assume C_{n+1} is the line $y = 0, -\infty < x < \infty$, and that $\theta(V)$ lies in $y > 0$. Then take $r(x)$ as the distance of $(x, 0)$ from the set

$$(V - C_{n+1}) \cup \sum_1^n H_p \cup D_{t_{n+1}} \cup D_{t_{n+1}+1}$$

and H_{n+1} as the curve $y = \frac{1}{2}r(x)$. Apply the inverse transformation to obtain the curve in the original plane.

The set of all H_p ($p = 1, 2, \cdots$) will then be completely defined by induction.

Properties (a) and (b) then hold, and also (c), (d) and (e), since these conditions are invariant under a homeomorphism. Further, by the construction, H_p separates C_p from each of $H_1, H_2, \cdots, H_{p-1}, D_{t_p}, D_{t_p+1}, C_0$ and from all $C_{p'}$ with $p' \neq p$. Thus (f), (h), (i), (j) hold and (*) $C_p \mid H_p \mid H_{p''}$ for $p'' < p$. If $p' > p$, then by (f) $C_{p'} \mid H_p \mid C_p$ and by (*) $C_{p'} \mid H_{p'} \mid H_p$. Hence, by Theorem I.17, $C_p \mid H_p \mid H_{p'}$. Thus (g) holds.

We next verify (k). $H_k \mid C_0 \mid D_{t_k+1}$ is impossible, by (b). If $C_0 \mid D_{t_k+1} \mid H_k$, then (i) and Axiom 3.3 give $C_0 \mid D_{t_k+1} \mid C_k$. This is impossible since $D_{t_k+1} \subset B(C_k)$. If $C_0 \mid H_k \mid D_{t_k+1}$, then this, with (i) and (j), contradicts Axiom 3.4. Hence $\mid C_0, H_k, D_{t_k+1} \mid^{\pm}$. This and (i) give by Axiom 3.2 $[C_0, H_k, D_{t_k+1}] \sim [C_0, H_k, D_{t_k+1}]$.

We now verify (l). If $H_k \mid H_{k'} \mid H_{k''}$, for example, then by (g) $C_{k'} \mid H_{k'} \mid H_{k''}$ and $C_{k'} \mid H_{k'} \mid H_k$. These three relations contradict Axiom 3.4 and similarly there is a contradiction if $H_{k'} \mid H_k \mid H_{k''}$ or $H_k \mid H_{k''} \mid H_{k'}$. Thus $\mid H_k, H_{k'}, H_{k''} \mid^{\pm}$. This and $C_k \mid H_k \mid H_{k'}$ give by Axiom 3.2 (1) $[H_k, H_{k'}, H_{k''}] \sim [C_k, H_{k'}, H_{k''}]$, with $\mid C_k, H_{k'}, H_{k''} \mid^{\pm}$. Similarly $C_{k'} \mid H_{k'} \mid H_{k''}$ gives (2) $[C_k, H_{k'}, H_{k''}] \sim [C_k, C_{k'}, H_{k''}]$, with $\mid C_k, C_{k'}, H_{k''} \mid^{\pm}$, and again $C_{k''} \mid H_{k''} \mid C_{k'}$ gives (3) $[C_k, C_{k'}, H_{k''}] \sim [C_k, C_{k'}, C_{k''}]$, with $\mid C_k, C_{k'}, C_{k''} \mid^{\pm}$. Combining (1), (2), and (3), we obtain $[H_k, H_{k'}, H_{k''}] \sim [C_k, C_{k'}, C_{k''}]$, and (l) is established. (m) and (n) are established in the same way, by means of (j) and (i) instead of (g). The proof of the theorem is thus complete.

The curves H_k enable us to isolate neighborhoods of the curve C_k. However, in order to ensure that these neighborhoods have a simple structure, we improve the choice of the H_k by means of the following theorems. We shall assume the H_k chosen fixed.

Definition. A curve γ_k will be said to *isolate* the curve C_k of V if it has the following properties:

(Is1) γ_k is an open curve tending to infinity in both directions;

(Is2) γ_k lies wholly in $\theta(V)$;

(Is3) $\gamma_k \cdot H_k = 0$ and $C_k \mid \gamma_k \mid H_k$;

(Is4) if $\mathfrak{D}(\gamma_k) \supset C_k$ and $\mathfrak{D}(C_k) \supset \gamma_k$, then the closed region $\omega_k = C_k \cup \gamma_k \cup \mathfrak{D}(\gamma_k) \cdot \mathfrak{D}(C_k)$ can be mapped o-homeomorphically on the region

$$W : 0 \leqq x \leqq (y^2 + 1)^{-1}, \ -\infty < y < \infty,$$

so that the inverse image of each line $x = \text{constant} > 0$ in W is on a curve D_t of $\theta(V)$, of $x = 0$ is C_k.

Theorem 17. *Let the curve γ_k isolate the curve C_k of V. Then in the region W of* (Is4) *γ_k has as image the curve $x = (y^2 + 1)^{-1}$. The point* (1, 0) *has as inverse a point P_k which divides γ_k into two extended cross-sections σ_k and μ_k. γ_k intersects precisely those curves D_t of $\theta(V)$ for t in an interval $[t_0, t_k)$ as in Theorem* I.41.

Proof. Since $x = 0$ has as inverse C_k, $x = (y^2 + 1)^{-1}$ must have as inverse the rest of the boundary of ω_k, i.e., γ_k. Let Q_k be the inverse of (0, 0), P_kQ_k the inverse of the straight line segment joining (0, 0) and (1, 0). Any subarc of P_kQ_k containing neither P_k nor Q_k is then a cross-section. If P_kQ_k itself met a curve of F twice, then, from Theorem I.8, the same would hold for a subarc not containing P_k or Q_k. This cannot arise, hence P_kQ_k is a cross-section. Thus the inverses of $x = (y^2 + 1)^{-1}$, $y \geqq 0$ and of $x = (y^2 + 1)^{-1}$, $y \leqq 0$ are extended cross-sections σ_k and μ_k. By Theorem I.41 and Corollary 2 to Theorem 14 P_kQ_k meets only those D_t of $\theta(V)$ of an interval $[t_0, t_k)$. Hence the same holds for σ_k and μ_k.

THEOREM 18. *Let C_k be a curve of V, Γ_0 an extended cross-section joining C_0 to ∞ in $\theta(V)$, and ϵ a given number > 0. A curve γ_k can be found so that*: (a) *it isolates C_k*; (b) $\gamma_k \cdot \Gamma_0 = 0$; (c) *the length of the interval $[t_0, t_k)$ of t for which γ_k meets D_t is less than ϵ.*

Proof. Suppose $D_{t_k} \subset B(C_k)$. Then draw a cross-section $P'Q_k$ joining P' on $D_{t'}$ in $\theta(V)$ to a point Q_k on C_k. By Theorem I.41, $t' < t_k$. By Theorem 14, a point P_0 on C_0 can be joined to $P_{t'}$ by a cross-section. By Theorem I.30, the set S formed by the curves D_t in $0 \leqq t < t_k$ plus C_k can be mapped o-homeomorphically on a strip $0 \leqq x \leqq 1$, $-\infty < y < \infty$, so that C_k becomes $x = 0$ and the curves D_t become the lines $x = t_k - t$.

Now H_k lies in $\theta(V)$. By (h) of Theorem 16, $C_k \mid H_k \mid D_{t_k}$. If $C_0 \mid D_{t_k} \mid H_k$, then $C_0 \mid D_{t_k} \mid C_k$, and this contradicts $D_{t_k} \subset B(C_k)$. Hence H_k lies in S and has image H'_k in the above strip. Let Γ'_0 be the image in the strip of the part of Γ_0 in S.

Let now $r(y)$ denote the distance of $(0, y)$ from the set $H'_k \cup \Gamma'_0$. Since Γ_0 tends to infinity, Γ'_0 can have no limit points on $x = 0$ and the same holds for H'_k. Hence $r(y) > 0$ and is continuous. Now set

$$\varphi(y_0) = \frac{\epsilon}{2r(0)(y_0^2 + 1)} \min_{0 \leqq y \leqq y_0} r(y) \qquad \text{for } y_0 \geqq 0,$$

$$\varphi(y_0) = \frac{\epsilon}{2r(0)(y_0^2 + 1)} \min_{y_0 \leqq y \leqq 0} r(y) \qquad \text{for } y_0 \leqq 0.$$

Then let γ'_k denote the curve given by $x = \varphi(y)$ and γ_k the inverse image of γ'_k. γ_k is then an open curve.

Now γ_k lies wholly in $\theta(V)$ and approaches the boundary of S in both directions. It cannot have limit points on any curve of $V - C_k$, for then C_k could be joined to points arbitrarily near C_0 or a $C_{k'}$ with $k' \neq k$ without crossing H_k. This is impossible by (f) and (j) of Theorem 16. By (h) we see that γ_k does not approach D_{t_k}. Hence γ_k tends to infinity in both directions.

No arc can be drawn connecting points on C_k and H_k and not crossing γ_k. For the image of such an arc would have to lie partly in the set $0 \leqq x < \varphi(y)$. But no point of $x = 0$ can be joined to H'_k in the strip without crossing γ'_k. Hence the image arc would have to tend to the boundary of $0 \leqq x \leqq \varphi(y)$. This implies, as in the preceding paragraph, that the inverse image tends to infinity in the plane, and we have reached a contradiction. Hence $C_k \mid \gamma_k \mid H_k$.

If now $\mathfrak{D}(\gamma_k) \supset C_k$ and $\mathfrak{D}(C_k) \supset \gamma_k$, then the region $\omega_k = C_k \cup \gamma_k \cup \mathfrak{D}(\gamma_k) \cdot \mathfrak{D}(C_k)$ has as image the set $0 \leqq x \leqq \varphi(y)$, $-\infty < y < \infty$. For the inverse image of the region $0 \leqq x \leqq \varphi(y)$ is a region whose boundary consists precisely of γ_k and C_k, which hence must coincide with ω_k.

An elementary transformation now transforms the region $0 \leqq x \leqq \varphi(y)$ onto the region $W: 0 \leqq x \leqq (y^2 + 1)^{-1}$ so that each line $x =$ constant becomes a line $x =$ constant and the curve $x = \varphi(y)$ becomes the curve $x = (y^2 + 1)^{-1}$.

Conditions (Is1), (Is2), (Is3), and (Is4) are now satisfied, hence γ_k isolates

C_k. Further γ_k does not meet Γ_0. Since $\varphi(y) < \epsilon$, γ_k' meets only lines $x = t$ for $| t - t_k | < \epsilon$, hence γ_k meets only curves D_t for $| t - t_k | < \epsilon$.

In case $D_{t_k} \not\subset B(C_k)$, we proceed in the same way, with D_{t_k+1} playing the rôle of C_0 and condition (i) of Theorem 16 replacing (h). Thus Theorem 18 is established.

We now apply the theorem, remarking that the inverse under T (of Theorem 15) of a line $x = \nu_k$, $0 \leqq y < \infty$, where ν_k is a constant with $| \nu_k | < 1$, is a curve Γ_0 of the type described. We then choose for each C_k in V an isolating curve γ_k not meeting the inverse of $x = \nu_k$. The precise choice of the constants ν_k will be indicated below.

2.4. Properties of the γ_k.

THEOREM 19. *The curves γ_k satisfy all the conditions of Theorem* 16, *if throughout H_k is replaced by γ_k.*

Proof. Since $C_k | \gamma_k | H_k$, if we choose $\mathfrak{D}(H_k)$ to include C_k, then $\gamma_k \subset \mathfrak{D}(H_k)$, while by Theorem 16, (f) $\cdots$ (j), $C_{k'}$, $H_{k'}$, D_{t_k}, D_{t_k+1}, C_0 all lie in $\mathfrak{D}^*(H)$. Hence γ_k fails to intersect any of these curves. Applying Theorem 17, we obtain (1) $C_k | \gamma_k | C_{k'}$, (2) $C_k | \gamma_k | H_{k'}$, (3) $C_k | \gamma_k | D_{t_k}$, (4) $C_k | \gamma_k | D_{t_k+1}$, (5) $C_k | \gamma_k | C_0$. From (2) and (f) we conclude $\gamma_k | H_{k'} | C_{k'}$. This and $H_{k'} | \gamma_{k'} | C_{k'}$ give that γ_k and $\gamma_{k'}$ do not intersect and further $C_k | \gamma_k | \gamma_{k'}$. It is now seen that (a) through (j) all hold with H_k replaced by γ_k. Since (k) $\cdots$ (n) are derived from (a) $\cdots$ (j), they must also hold with H_k replaced by γ_k. Thus Theorem 19 holds.

We shall refer to Theorem 19 (a), (b), $\cdots$ as meaning Theorem 16 (a), (b), $\cdots$ with H_k replaced by the isolating curve γ_k.

THEOREM 20. *If the subcurves σ_k and μ_k are properly named, then the image of γ_k under the homeomorphism T of Theorem* 15 *is a curve $x = g_k(y)$ defined in the interval $[t_{0k}, t_k)$, whereby $g_k(y)$ is a two-valued function $(s_k(y), m_k(y))$ with $\sigma_k \leftrightarrow x = s_k(y)$, $\mu_k \leftrightarrow x = m_k(y)$ and*

$$-1 < s_k(t_{0k}) = m_k(t_{0k}) < 1,$$

$$-1 < s_k(y) < m_k(y) < 1 \textit{ for } y \textit{ in } (t_{0k}, t_k),$$

$$\lim_{y \to t_k} s_k(y) = \lim_{y \to t_k} m_k(y) = k/| k |,$$

$$| \nu_k | < | g_k(y) | < 1,$$

***and** $s_k(y)$ and $m_k(y)$ are continuous in the interval $[t_{0k}, t_k)$.*

Proof. Since $\gamma_k \subset \theta(V)$ and by Theorem 17, it follows that σ_k and μ_k have as images curves $x = s_k(y)$ and $x = m_k(y)$ defined and continuous in the interval $[t_{0k}, t_k)$ and meeting only for $y = t_{0k}$. Hence, if σ_k and μ_k are properly named,

$s_k(y) < m_k(y)$ for $y \neq t_{0k}$. Also $\lim_{y \to t_k} g_k(y) = \pm 1$, since γ_k tends to infinity in both directions. Since γ_k does not meet the inverse of $x = \nu_k$, we must have $\lim_{y \to t_k} s_k(y) = \lim_{y \to t_k} m_k(y) = \pm 1$ and $|\nu_k| < |g_k(y)| < 1$.

Suppose $k > 0$, whence $C_k \subset V^+$, $|C_0, C_k, D_{t_k+1}|^+$ and, by Theorem 19 (k), $|C_0, \gamma_k, D_{t_k+1}|^+$. If $\lim_{y \to t_k} g_k(y) = -1$, then a triangle $M_0 M_k M_{t_k+1}$ through points M_0: $x = \nu_k$, $y = 0$, M_k : the point with minimum y-coördinate on $x = g_k(y)$, and M_{t_k+1}: $x = \nu_k$, $y = t_k + 1$ has negative orientation. Since T is an o-homeomorphism, this implies $|C_0, \gamma_k, D_{t_k+1}|^-$, and a contradiction is reached. Hence $\lim_{y \to t_k} g_k(y) = 1 = k/|k|$. Similarly, if $k < 0$, $\lim_{y \to t_k} g_k(y) = -1 = k/|k|$. The theorem is thus established.

THEOREM 21. (a) $\omega_k - C_k \subset \theta(V)$. (b) $\omega_k \cdot \omega_{k'} = 0$ *for* $k \neq k'$. (c) *The image of* $\omega_k - C_k$ *under* T *is the set* $s_k(y) \leqq x \leqq m_k(y)$ *for* y *in* $[t_{0k}, t_k)$.

Proof. Since $C_0 | \gamma_k | C_k$, if $\mathfrak{D}(C_0) \supset C_k$, then $\mathfrak{D}(C_0) \supset \mathfrak{D}(\gamma_k) \supset \mathfrak{D}^*(C_k)$. Hence $\omega_k \subset \mathfrak{D}(C_0)$. Similarly $\omega_k \subset \mathfrak{D}(C_{k'})$ for $k' \neq k$. Also, by definition, $\omega_k - C_k \subset \mathfrak{D}(C_k)$. Hence $\omega_k - C_k \subset \theta(V)$, and this gives (a). (b) follows from Theorem 19 (f) and (g) by a similar reasoning.

From (Is4) it follows that the image of $\omega_k - C_k$ under T is a set bounded by $x = g_k(y)$ and including the line segments y = constant joining $x = s_k(y)$ and $x = m_k(y)$. Hence $T(\omega_k - C_k)$ is the region $s_k(y) \leqq x \leqq m_k(y)$, y in $[t_{0k}, t_k)$. This gives (c).

2.5. **Adjustment of the transformations T^1, T^2.** The transformations T^1, T^2 of Theorem 15 (see the convention on superscripts in the fourth paragraph of §2.2) are not in general such that the two curves $x = g_k^1(y)$ and $x = g_k^2(y)$ coincide for each k. In the present section we shall indicate how the choices of both the γ_k^1 and γ_k^2 and T^1 and T^2 can be fixed in such a way that $x = g_k^1(y)$ and $x = g_k^2(y)$ coincide.

We consider first the case $k > 0$. We then specify the values of the constants ν_k, which will be the same for F_1 and F_2. If C_1 is defined, we take ν_1 so that $\frac{1}{2} < \nu_1 < 1$, and then γ_1^1 and γ_1^2 under the restrictions first that γ_1^1 does not meet the inverse under T^1 of the line $x = \nu_1$ and that γ_1^2 does not meet the inverse under T^2 of $x = \nu_1$, secondly that the corresponding intervals $[t_{01}^1, t_1^1)$ and $[t_{01}^2, t_1^2)$ are the same. This latter is possible by virtue of the fact that $t_1^1 = t_1^2 = t_1$ by the isomorphism of F_1 and F_2, that $t_{01}^1 < t_1$ or $t_{01}^1 > t_1$ according as $t_{01}^2 < t_1$ or $t_{01}^2 > t_1$ respectively by the isomorphism and Corollary 2 to Theorem 14, and finally that t_{01}^1 and t_{01}^2 can be taken arbitrarily close to t_1 by Theorem 18. Thus we can write $t_{01}^1 = t_{01}^2 = t_{01}$.

Next (if C_2 is defined) we choose a constant ν_2, $\frac{3}{4} < \nu_2 < 1$, and further larger than both $g_1^1(t_{01})$ and $g_1^2(t_{01})$. Then choose γ_2^1 and γ_2^2 so that they do not meet the respective inverses of $x = \nu_2$, $0 \leqq y < \infty$, further so that $t_{02}^1 = t_{02}^2 = t_{02}$, $t_2^1 = t_2^2 = t_2$, and also so that $t_{02} \neq t_{01}$.

In general, if we have defined $\gamma_1, \cdots, \gamma_p$, then we choose ν_{p+1} so that $1 - 2^{-p-1} < \nu_{p+1} < 1$ and so that $\nu_{p+1} > g_p^1(t_{0p})$, $\nu_{p+1} > g_p^2(t_{0p})$. Then choose γ_{p+1}^1 and γ_{p+1}^2 not to meet the inverses of $x = \nu_{p+1}$, $0 \leqq y < \infty$ and so that $t_{0,p+1}^1 = t_{0,p+1}^2 = t_{0,p+1}$, $t_{p+1}^1 = t_{p+1}^2 = t_{p+1}$, while $t_{0,p+1} \neq t_{0k}$ for $k = 1, \cdots, p$.

With these choices fixed, we then see that $x = g_k(y)$ lies to the right of $x = \nu_k$, while the line $x = \nu_{k+1}$ lies to the right of the point P_k: $y = t_{0k}$, $x = g_k(t_{0k})$. Further $\lim_{k \to \infty} \nu_k = 1$, $\nu_{k+1} > \nu_k$, and $t_{0k} \neq t_{0k'}$ for $k \neq k'$. These relations hold for all $k = 1, 2, \cdots$.

For $k < 0$ we proceed in a similar way, choosing the constants $\nu_{-1}, \nu_{-2}, \cdots$ and the curves $\gamma_{-1}^1, \gamma_{-2}^1, \cdots, \gamma_{-1}^2, \gamma_{-2}^2, \cdots$, so that $-1 + (1/2^k) > \nu_k > -1$, that $\nu_{k-1} < g_k^1(t_{0k})$ and $\nu_{k-1} < g_k^2(t_{0k})$, and that $x = g_k(y)$ lies to the left of $x = \nu_k$, while $\lim_{k \to -\infty} \nu_k = -1$, $\nu_{k-1} < \nu_k$, $t_{0k}^1 = t_{0k}^2 = t_{0k}$, $t_k^1 = t_k^2$ but $t_{0k} \neq t_{0k'}$ for $k \neq k'$.

THEOREM 22. *If $k' \neq k''$, $k'/|k'| = k''/|k''|$, and y is such that both $g_{k'}^1(y)$ and $g_{k''}^1(y)$ are defined, then either*

$$\begin{cases} s_{k'}^1(y) \leqq m_{k'}^1(y) < s_{k''}^1(y) \leqq m_{k''}^1(y), \\ s_{k'}^2(y) \leqq m_{k'}^2(y) < s_{k''}^2(y) \leqq m_{k''}^2(y) \end{cases}$$

or else

$$\begin{cases} s_{k''}^1(y) \leqq m_{k''}^1(y) < s_{k'}^1(y) \leqq m_{k'}^1(y), \\ s_{k''}^2(y) \leqq m_{k''}^2(y) < s_{k'}^2(y) \leqq m_{k'}^2(y). \end{cases}$$

Proof. Case I. $k > 0$. We can suppose (without restriction) that $k' < k''$. Suppose first that $t_{k'} \neq t_{k''}$. The two intervals $[t_{0k'}, t_{k'})$ and $[t_{0k''}, t_{k''})$ must overlap. If $t_{0k'}$ is contained in $[t_{0k''}, t_{k''})$, then, by the above construction, $g_{k''}^1(t_{0k'}) > g_{k'}^1(t_{0k'})$. By continuity this must hold for all y in the common interval, since the curves do not intersect. If $t_{k'}$ is contained in $[t_{0k''}, t_{k''})$, then, as $y \to t_{k'}$, $g_{k'}^1(y) \to 1$, while $g_{k''}^1(y) \to g_{k''}^1(t_{k'}) < 1$. Hence near $t_{k'}$ $g_{k'}^1(y) > g_{k''}^1(y)$ and from continuity this holds throughout. Next, if one but not both of $t_{0k''}$, $t_{k''}$ is contained in $[t_{0k'}, t_{k'})$, then either $t_{k'}$ or $t_{0k'}$ is in $[t_{0k''}, t_{k''})$. This returns us to the previous case. If both $t_{0k''}$ and $t_{k''}$ are in $[t_{0k'}, t_{k'})$, then near $t_{k''}$ $g_{k''}^1(y) > g_{k'}^1(y)$ and by continuity this holds throughout. Thus in all these cases the order of the points is determined purely by the intervals involved and the fact that $k' < k''$. It follows that the points $x = g_{k'}^2(y)$ and $x = g_{k''}^2(y)$ lie in the same order, for any y in the common domain of definition of $g_{k'}^2(y)$ and $g_{k''}^2(y)$ (which is the same as the common domain of $g_{k'}^1(y)$ and $g_{k''}^1(y)$).

Next suppose $t_{k'} = t_{k''}$. The only possible cases are $t_{0k'} < t_{k'}$, $t_{0k''} < t_{k''}$ or else $t_{0k'} > t_{k'}$, $t_{0k''} > t_{k''}$. Consider the first of these cases. Suppose $t_{0k'} < t_{0k''} < t_{k'}$. At $t_{0k''}$ we must then have either $g_{k''}^1(t_{0k''}) < s_{k'}^1(t_{0k''}) \leqq m_{k'}^1(t_{0k''})$ or $g_{k''}^1(t_{0k''}) > m_{k'}^1(t_{0k''}) \geqq s_{k'}^1(t_{0k''})$. For if $s_{k'}^1(t_{0k''}) < g_{k''}^1(t_{0k''}) < m_{k'}^1(t_{0k''})$, then by Theorem 21 (c) $\omega_{k'}$ contains $\gamma_{k''}$. This is impossible by the fact that $\omega_{k'} \cdot \omega_{k''} = 0$, by Theorem 21 (a).

If $g_{k''}^1(t_{0k''}) < g_{k'}^1(t_{0k''})$, then it is possible to find a closed curve through the

points $(g^1_{k'}(t_{0k'}), 0)$, $(g^1_{k'}(t_{0k'}), t_{0k'})$, $(g^1_{k''}(t_{0k''}), t_{0k''})$, meeting the curves $y = 0$, $x = g^1_{k'}(y)$, and $x = g^1_{k''}(y)$ only at these three points, and lying wholly in the half-strip $|x| < 1, 0 \leqq y < \infty$. Moreover, such a curve has positive orientation. Applying the inverse of T^1 and the definition of the chordal relations, we obtain that $| C^1_0, \gamma^1_{k'}, \gamma^1_{k''} |^+$. By Theorem 19 (n) $[C^1_0, \gamma^1_{k'}, \gamma^1_{k''}] \sim [C^1_0, C^1_{k'}, C^1_{k''}]$. It follows that we can have $g^1_{k''}(t_{0k''}) < g^1_{k'}(t_{0k''})$ only if $| C^1_0, C^1_{k'}, C^1_{k''} |^+$. Similarly we find $g^1_{k''}(t_{0k''}) > g^1_{k'}(t_{0k''})$ only if $| C^1_0, C^1_{k'}, C^1_{k''} |^-$. It thus follows from the isomorphism that $g^2_{k''}(t_{0k''}) < g^2_{k'}(t_{0k''})$ according as $g^1_{k''}(t_{0k''}) < g^1_{k'}(t_{0k''})$. By continuity the same condition holds for any y in the common interval. Thus, under the assumption $t_{0k'} < t_{0k''} < t_{k'}$, the desired inequalities hold. Exactly the same type of discussion holds if $t_{0k''} < t_{0k'} < t_{k'}$.

A similar discussion holds in case $t_{0k'} > t_{k'}$ and $t_{0k''} > t_{k''}$. Thus in all cases under Case I the inequalities hold.

Case II, $k < 0$. The same reasoning holds, and thus the theorem follows.

THEOREM 23. *T^1 can be so chosen that it satisfies Theorem* 15 *and further that $g^1_k(y)$ coincides with $g^2_k(y)$ for all k.*

Proof. Case I, $k > 0$. We shall first make certain extensions in the domain of definition of the curves $g^1_k(y)$ and $g^2_k(y)$ so that the inequalities of Theorem 22 continue to hold, with the exception of a possible equality instead of inequality in certain cases.

We make these changes in two stages. First (A) suppose $g^1_1(y)$ has interval of definition $t_{01} \leqq y < t_1$. It may happen that (A1) some other curve $x = g^1_k(y)$ crosses the line $y = t_{01}$ between $x = g^1_1(t_{01})$ and $x = 1$. (It will then actually cross the line, since $t_{0k} \neq t_{01}$.) If so, we set $\delta_1 = 0$ and leave $g^1_1(y)$ unchanged. The same situation then holds for $g^2_1(y)$, by Theorem 22, and again there is no change. If this condition fails to hold, then there are two possibilities: either (A2) for δ sufficiently small no curve $g^1_k(y)$ $(k > 1)$ meets the rectangle $t_{01} - \delta \leqq y \leqq t_{01}$, $g^1_1(t_{01}) \leqq x \leqq 1$, or (A3) no such δ can be found. In the first of these cases we leave the curves $g^1_1(y)$ and $g^2_1(y)$ unchanged at this stage and take $\delta_1 = 0$.

In case (A3) we choose $\delta_1 > 0$ so small that the following properties hold:

(a) no other curve $g^1_k(y)$ meets the line segment $x = g^1_1(t_{01})$, $t_{01} - \delta_1 \leqq y \leqq t_{01}$;

(b) no other curve $g^2_k(y)$ meets the line segment $x = g^2_1(t_{01})$, $t_{01} - \delta_1 \leqq y \leqq t_{01}$;

(c) there is a curve $g^1_{k_1}(y)$ crossing the line $y = t_{01} - \delta_1$ in the interval $g^1_1(t_{01}) < x < 1$ and no other curve $g^1_k(y)$ crosses the line in the interval $g^1_1(t_{01}) < x < g^1_{k_1}(y)$;

(d) for the same k_1, $g^2_{k_1}(y)$ has the analogous property for $g^2_1(y)$.

In order to find this δ_1, we first restrict δ_1 to be so small that at least (a) and (b) hold. This is possible since the curves $g^1_k(y)$ and $g^2_k(y)$ tend uniformly to the boundary. Next let $\varphi^1_1(y)$ be the $g^1_{k_0}(y)$ crossing the line $y = t_{01}$, $0 < x < g^1_1(t_{01})$ with maximal value of x, or else, if there is no such $g^1_{k_0}(y)$, set $\varphi^1_1(y) \equiv 0$. $\varphi^2_1(y)$ will be determined in the same way, and the index k_0 will be the same in both cases, by Theorem 22. We now restrict δ_1 further to be so small that

$t_{01} - \delta_1 \leqq y \leqq t_{01}$ lies within the interval of definition of $\varphi_1^1(y)$ and further so small that no curve $g_k^1(y)$ crosses the region $\varphi_1^1(y) \leqq x \leqq g_1^1(t_{01})$, $t_{01} - \delta_1 \leqq y \leqq t_{01}$ and that no other $g_k^2(y)$ crosses the region $\varphi_1^2(y) \leqq x \leqq g_1^2(t_{01})$, $t_{01} - \delta_1 \leqq y \leqq t_{01}$. Finally, since infinitely many curves $g_k^1(y)$ meet the rectangle $t_{01} - \delta_1 \leqq y \leqq t_{01}$, $g_1^1(t_{01}) \leqq x \leqq 1$, we can take δ_1 further so small that a $g_{k_1}^1(y)$ crosses the line $y = t_{01} - \delta_1$, $g_1^1(t_{01}) \leqq x \leqq 1$, with $t_{01} - \delta_1 \neq t_{0k_1}$, and that no other $g_k^1(y)$ crosses $y = t_{01} - \delta_1$ between $x = g_1^1(t_{01})$ and $x = g_{k_1}^1(t_{01})$. It follows from Theorem 22 that, with the same δ_1 and k_1, the function $g_{k_1}^2(y)$ has the same properties relative to $g_1^2(y)$. The conditions (a), (b), (c), (d) are now satisfied. We now extend the definition of $g_1^1(y)$ by setting $g_1^1(y) = g_1^1(t_{01})$ for $t_{01} - \delta_1 \leqq y \leqq t_{01}$, and similarly $g_1^2(y) = g_1^2(t_{01})$ in the same interval. Further we set $s_1^1(y) = m_1^1(y) = g_1^1(y)$, $s_1^2(y) = m_1^2(y) = g_1^2(y)$ in this interval.

In case (B) $t_1 < t_{01}$, then we have the analogous discussion for cases (B1), (B2), (B3).

Now we extend this process by induction to all curves $g_k^1(y)$ and $g_k^2(y)$. Thus, if it has been carried out for $k = 1, 2, \cdots, r - 1$, then we consider $g_r^1(y)$ and $g_r^2(y)$. The discussion is then exactly as with $g_1^1(y)$ and $g_1^2(y)$, except that in each reference to the other curves $g_k^1(y)$ and $g_k^2(y)$ for $k < r$ the curves as extended under (A3) or (B3) will be meant.

As a result of this, we obtain a new family of curves $g_k^1(y)$ and $g_k^2(y)$ whose definition intervals are $[t_{0k} \pm \delta_k, t_k)$, with $\delta_k \geqq 0$, so that no two curves intersect and the inequalities of Theorem 22 still hold. Finally, as a result of those changes under (A3) and (B3), for the new end-points $\bar{t}_{0k} = t_{0k} \pm \delta_k$ the only possible cases are (A1) and (A2), (B1) and (B2).

We now carry the extensions one stage further. All references will be to the curves as they stand after the above extensions. For $g_1^1(y)$ in case (A1) we choose $g_{k_1}^1(y)$ as the curve $g_k^1(y)$ crossing $y = \bar{t}_{01}$ between $x = g_1^1(\bar{t}_{01})$ and $x = 1$ with minimum value of x. $g_{k_1}^2(y)$ will have the same property relative to $g_1^2(y)$. We next choose $\delta > 0$ so small that no other $g_k^1(y)$ crosses the region $\bar{t}_{01} - \delta \leqq y \leqq \bar{t}_{01}$, $g_1^1(\bar{t}_{01}) \leqq x \leqq g_{k_1}^1(y)$ and no other $g_k^2(y)$ crosses the region $\bar{t}_{01} - \delta \leqq y \leqq \bar{t}_{01}$, $g_1^2(\bar{t}_{01}) \leqq x \leqq g_{k_1}^2(y)$, δ being also taken so small that $g_{k_1}^1(y)$ is defined in $\bar{t}_{01} - \delta \leqq y \leqq \bar{t}_{01}$. We then extend $g_1^1(y)$ and $g_1^2(y)$ by setting

$$s_1^1(y) = m_1^1(y) = g_1^1(y) = \frac{\bar{t}_{01} - y}{\delta} s_{k_1}^1(y) + \left(1 - \frac{\bar{t}_{01} - y}{\delta}\right) g_1^1(\bar{t}_{01}),$$

$$s_1^2(y) = m_1^2(y) = g_1^2(y) = \frac{\bar{t}_{01} - y}{\delta} s_{k_1}^2(y) + \left(1 - \frac{\bar{t}_{01} - y}{\delta}\right) g_1^2(\bar{t}_{01})$$

for $\bar{t}_{01} - \delta \leqq y \leqq \bar{t}_{01}$. The extensions to the functions lie within the above regions and on the extensions $g_1^1(\bar{t}_{01}) \leqq g_1^1(y) < g_{k_1}^1(y)$ except for $y = \bar{t}_{01} - \delta$, when $g_1^1(y) = g_{k_1}^1(y)$. Similarly $g_1^2(y) < g_{k_1}^2(y)$ except that $g_1^2(\bar{t}_{01} - \delta) = g_{k_1}^2(\bar{t}_{01} - \delta)$. The inequalities of Theorem 22 thus continue to hold, with a new inequality at one point, namely, at the new end-point of the interval of definition of $g_1^1(y)$ and $g_1^2(y)$.

In case (A2) we proceed similarly, replacing $g^1_{k_1}(y)$ by the function $\psi_1(y) \equiv 1$. For the new functions $g^1_1(y)$ and $g^2_1(y)$ we then have $g^1_1(\bar{t}_{01} - \delta) = g^2_1(\bar{t}_{01} - \delta) = 1$. Since no curve $g^1_k(y)$ (for $k > 1$) crosses the rectangle $\bar{t}_{01} - \delta \leqq y \leqq \bar{t}_{01}$, $g^1_1(\bar{t}_{01}) \leqq x \leqq 1$, and similarly no $g^2_k(y)$ crosses the rectangle $\bar{t}_{01} - \delta \leqq y \leqq \bar{t}_{01}$, $g^2_1(\bar{t}_{01}) \leqq x \leqq 1$, we conclude that the inequalities of Theorem 22 still hold.

In cases (B1) and (B2) we proceed in the same way.

We now extend this process by induction to all curves $g^1_k(y)$ and $g^2_k(y)$ for $k > 0$, as above in cases (A3) and (B3), using at each stage the already extended curves for the lower indices. Inasmuch as the values $x = g^1_k(\bar{t}_{0k}) = g^1_k(t_{0k})$ tend to 1 as $k \to \infty$, and on the extensions to the curves $x \geqq g^1_k(\bar{t}_{0k})$, the exten-

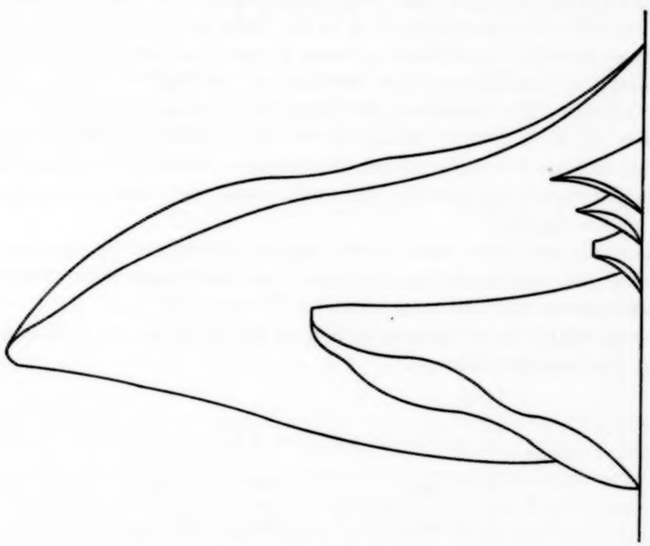

FIG. 2

sions (if infinite in number) tend uniformly to the boundary $x = 1$. Further the order relations of Theorem 22 will continue to hold. (See Figure 2.)

Now as a result of these extensions the set of curves $x = g^1_k(y)$ forms a set of curve arcs, each of which joins either a point on an arc to a point on another or a point on the line $x = 1$ to a point on an arc, and there are no free end-points in $x < 1$. Further the intersection points of the arcs are either countably infinite or finite, and, if infinite, tend to $x = 1$. At each intersection point at most a finite number of arcs meet. It follows that these arcs divide the region $0 < x < 1, 0 \leqq y < \infty$ into an at most countably infinite number of connected regions. The same holds for the curves $x = g^2_k(y)$.

The interior of each region determined by the $x = g^1_k(y)$ has the structure $\varphi(y) < x < \psi(y)$, where $\varphi(y)$ and $\psi(y)$ are continuous functions in the same

finite or infinite y-interval. That $\varphi(y)$ is continuous and well-defined follows since on each of the curve arcs x is a continuous function of y. To determine $\varphi(y)$ we take any point (x_0, y_0) of the region under consideration and let $(\varphi(y_0), y_0)$ be the boundary point to the left of (x_0, y_0) on the line $y = y_0$. If $\varphi(y_0) = 0$, then $\varphi(y) \equiv 0$. If $\varphi(y_0) \neq 0$, then the rest of the boundary $x = \varphi(y)$ can be obtained by following along the curve arc through $(\varphi(y_0), y_0)$ letting y increase and decrease indefinitely. If a multiple point is reached, then the continuation arc is chosen to have maximum possible value of x. If, as y approaches a value y_1, $\varphi(y)$ approaches 1, then $\varphi(y)$ ceases to be defined beyond y_1 (above or below, according to the case).

$\psi(y_0)$ can be defined (in the same interval as $\varphi(y)$) as the function equal to 1 if there is no curve arc meeting $y = y_0$ to the right of (x_0, y_0) and as equal to the minimum possible x-coördinate of such a curve arc when it exists. $\psi(y)$ is then necessarily continuous when less than 1. If $\psi(y_0) = 1$, then $\psi(y)$ remains continuous. For otherwise the curves $x = g_k^1(y)$ would have a limit point on $y = y_0$, $x_0 < x < 1$. Since the $x = g_k^1(y)$ tend uniformly to 1, this is possible only if some $x = g_k^1(y)$ crosses the segment, contrary to assumption.

Since the functions $\varphi(y)$ and $\psi(y)$ are continuous, they must form the whole boundary of the region.

Now the same discussion holds for the regions determined by the $x = g_k^2(y)$, and, moreover, the regions are in one-to-one correspondence with those of the $g_k^1(y)$ in accordance with the inequalities of Theorem 22.

We can now obtain an o-homeomorphism of the set $0 \leqq x \leqq 1, 0 \leqq y < \infty$ on itself by the transformation:

$$\bar{T}: \begin{cases} \bar{y} = y; \\ \bar{x} = 0 \text{ for } x = 0, \ \bar{x} = 1 \text{ for } x = 1; \\ \bar{x} = s_k^2(y) \qquad \text{for } x = s_k^1(y); \\ \bar{x} = m_k^2(y) \qquad \text{for } x = m_k^1(y); \end{cases}$$

with $\bar{x}$ varying linearly with x between these values. This transformation leaves each line $y = y_0$ invariant, is also monotone increasing and continuous in x for fixed $y = y_0$, hence gives a homeomorphism of each line $y = y_0$ onto itself, and is therefore one-to-one everywhere. Further it takes each region $\varphi^1(y) < x < \psi^1(y)$ determined by the $g_k^1(y)$ onto the corresponding one $\varphi^2(y) < x < \psi^2(y)$ determined by the $g_k^2(y)$. Since the $\varphi^1(y)$ and $\psi^1(y)$ are continuous functions, and $\bar{T}$ varies linearly between $\varphi^1(y)$ and $\psi^1(y)$, $\bar{T}$ is a homeomorphism in each region plus boundary. $\bar{T}$ is thus a homeomorphism on the region $0 \leqq x < 1, 0 \leqq y < \infty$, and, since $\bar{y} = y$, also on the boundary $x = 1$. $\bar{T}$ is also an o-homeomorphism.

In a similar manner in Case II, $k < 0$, an o-homeomorphism $\bar{T}$ of the region $-1 \leqq x \leqq 0, 0 \leqq y < \infty$ on itself can be defined leaving the line $x = 0$ and each line $y = y_0$ invariant and transforming all the $x = g_k^1(y)$ for $k < 0$ on the corresponding $x = g_k^2(y)$. The resulting transformation $\bar{T}$ on the whole half-

strip $-1 \leqq x \leqq 1$, $0 \leqq y < \infty$ is then single-valued and an o-homeomorphism of the set on itself.

The transformation $\bar{T}T^1$ can now be chosen to replace the original choice of T^1, and Theorem 23 is established.

2.6. **Map of F_1 on F_2.** We assume the transformations T of Theorem 15 and the γ_k are now fixed so that Theorem 23 holds.

LEMMA. *Let W be a homeomorphism of C_0 onto the line $y = 0$, $-1 < x < 1$ such that the homeomorphism $T^{-1}W$ of C_0 on itself preserves orientation on C_0. Then there exists an o-homeomorphism $\tilde{T}$ of $C_0 \cup \theta(V)$ on the strip $-1 < x < 1$, $0 \leqq y < \infty$ such that $\tilde{T} = T$ along each γ_k, $\tilde{T} = W$ along C_0 and the image of each curve D_t is the line $y = t$.*

Proof. By Theorem 20, $|\nu_k| < |g_k(y)| < 1$. Also $|1 - (1/2^{|k|})| < |\nu_k|$. Hence the distance $r(x)$ of a point $(x, 0)$ on $y = 0$, $-1 < x < 1$ from the point set formed by all the curves $x = g_k(y)$ is greater than 0. The function $y = \psi(x)$ in $-1 \leqq x \leqq 1$ defined by

$$\psi(x_0) = \frac{1 - x_0^2}{2} \min_{0 \leqq x \leqq x_0} r(x), \qquad x_0 \geqq 0;$$

$$\psi(x_0) = \frac{1 - x_0^2}{2} \min_{x_0 \leqq x \leqq 0} r(x), \qquad x_0 \leqq 0,$$

is continuous, monotone strictly decreasing for $x \geqq 0$, monotone strictly increasing for $x \leqq 0$, and no curve $x = g_k(y)$ intersects the set $0 \leqq y \leqq \psi(x)$, $-1 \leqq x \leqq 1$.

Since $T^{-1}W$ preserves orientation on C_0, the "transposed" transformation $T(T^{-1}W)T^{-1} = WT^{-1}$ of $-1 < x < 1$, $y = 0$ onto itself must also preserve orientation and can be extended continuously to the end-points, which it leaves fixed. Such a transformation can be extended to the set $0 \leqq y \leqq \psi(x)$, $-1 \leqq x \leqq 1$ so as to leave each point of the curve $y = \psi(x)$ fixed and in general to leave y invariant. We write the extended homeomorphism as

$$T^0\colon\ x^0 = \varphi(x, y), \qquad y^0 = y.$$

If we define T^0 as the identity outside of this set, then T^0 becomes an o-homeomorphism of $-1 \leqq x \leqq 1$, $0 \leqq y < \infty$ on itself. The product $\tilde{T} = T^0T$ then coincides with T along each γ_k, $\tilde{T}(D_t)$ is the line $y = t$ and along C_0, $\tilde{T} = T^0T = WT^{-1}T = W$.

THEOREM 24. *There is an o-homeomorphism of each set ω_k^1 on the corresponding set ω_k^2 such that the image of each arc of a curve D_t^1 in ω_k^1 is an arc of the curve D_t^2 and that the image of γ_k^1 is γ_k^2, of C_k^1 is C_k^2, of σ_k^1 is σ_k^2, of μ_k^1 is μ_k^2.*

Proof. By condition (Is4) and Theorem 17, ω_k^1 and ω_k^2 can both be mapped o-homeomorphically on the set W: $0 \leqq x \leqq (y^2 + 1)^{-1}$, $-\infty < y < \infty$ so that each line $x = \text{constant} > 0$ in W is an image of an arc of a D_t, of $x = 0$ is C_k.

Further, by construction, the intervals $[t_{0k}, t_k)$ are the same for γ_k^1 and γ_k^2. Hence by suitable choice of the o-homeomorphisms of ω_k^1 and ω_k^2 on W we can obtain that each arc D_t^1 in ω_k^1 has the same image x = constant in W as the arc D_t^2. If we show further that σ_k^1 and σ_k^2, μ_k^1 and μ_k^2 have the same images, then the theorem will be established.

Now σ_k^1 is defined as the inverse of $x = s_k^1(y)$, μ_k^1 of $x = m_k^1(y)$, σ_k^2 of $x = s_k^2(y)$, μ_k^2 of $x = m_k^2(y)$ (without the extensions to these functions temporarily introduced in the proof of Theorem 23). From Theorem 21 it follows that the set $s_k^1(y) \leqq x \leqq m_k^1(y)$ can be mapped o-homeomorphically on W minus the line $x = 0$ so that $x = g_k^1(y)$ becomes the curve $x = 1/(y^2 + 1)$, and so that each line $y = t$ becomes a line x = constant. The set $s_k^2(y) \leqq x \leqq m_k(y)$ can similarly be mapped on W minus the line $x = 0$. Further, $s_k^1(y) \equiv s_k^2(y)$, $m_k^1(y) \equiv m_k^2(y)$, by Theorem 23. Since both of these above homeomorphisms preserve orientation, we conclude that $s_k^1(y)$ and $s_k^2(y)$ have the same image, i.e., either $x = 1/(y^2 + 1)$, $y \geqq 0$ or $x = 1/(y^2 + 1)$, $y \leqq 0$, and that $m_k^1(y)$ and $m_k^2(y)$ have the same image. Hence σ_k^1 and σ_k^2 have the same image, μ_k^1 and μ_k^2 have the same image. The theorem is now established.

THEOREM 25. *There is an o-homeomorphism of $\lambda(V^1)$ onto $\lambda(V^2)$ taking each C_0^1 onto C_0^2, C_k^1 onto C_k^2, D_t^1 onto D_t^2.*

Proof. Under T^1, the set $(\lambda(V^1) - \sum_k \omega_k^1) \cup \sum_k \gamma_k^1$ is mapped o-homeomorphically on the strip $-1 < x < 1, 0 \leqq y < \infty$ minus the sets $s_k(y) < x < m_k(y)$. By Theorem 23, under T^2, $(\lambda(V^2) - \sum_k \omega_k^2) \cup \sum_k \gamma_k^2$ is mapped o-homeomorphically on the same set. In both cases the curve D_t has as image the line $y = t$. It follows that $(T^2)^{-1}T^1$ takes $(\lambda(V^1) - \sum \omega_k^1) \cup \sum \gamma_k^1$ o-homeomorphically onto $(\lambda(V^2) - \sum \omega_k^2) \cup \sum \gamma_k^2$ and takes D_t^1 onto D_t^2. Further the image of each γ_k^1 is γ_k^2, and, furthermore, of σ_k^1 is σ_k^2, of μ_k^1 is μ_k^2.

By Theorem 25, there is an o-homeomorphism of each ω_k^1 onto ω_k^2. Moreover, this homeomorphism coincides with $(T^2)^{-1}T^1$ along σ_k^1, since σ_k^1 meets each D_t^1 at most once, at a point P_t^1, and each point P_t^1 on σ_k^1 has image P_t^2 on D_t^2 under both transformations. Thus they coincide along σ_k^1, and similarly along μ_k^1, hence on γ_k^1. But γ_k^1 is the boundary of ω_k^1 in $\theta(V^1)$. Hence by Theorem 21(b) the transformation $(T^2)^{-1}T^1$ can be extended to be single-valued over all the sets $\omega_k^1 \cup C_k^1$, becomes an o-homeomorphism of $\lambda(V^1)$ onto $\lambda(V^2)$ with the desired properties.

THEOREM 26. *F_1 can be mapped o-homeomorphically on F_2 by a transformation T such that $T(C^1) = C^2 = f(C^1)$ for each curve C^1 of F_1.*

Proof. By Theorem 25, there is an o-homeomorphism T_1 of $\lambda(V_1^1)$ onto $\lambda(V_1^2)$ preserving curves of the family. There is a similar transformation $T_{1,k}$ of each $\lambda(V_{1,k}^1)$ onto the corresponding $\lambda(V_{1,k}^2)$. $\lambda(V_1^1)$ and $\lambda(V_{1,k}^1)$ overlap only along the curve $C_{1,k}^1$. Hence $T_{1,k}^{-1} \cdot T_1$ defines a homeomorphism of $C_{1,k}^1$ onto itself. Since $C_{1,k}^1$ is an open curve tending to infinity in both directions, and since $T_{1,k}$

and T_1 both are o-homeomorphisms, $T_{1,k}^{-1}T_1$ must leave orientation invariant on $C^1_{1,k}$.

$T_{1,k}$ is defined by means of an o-homeomorphism $T^1_{1,k}$ of $C^1_{1,k} \cup \theta(V^1_{1,k})$ onto the half-strip $-1 < x < 1$, $0 \leqq y < \infty$ and an o-homeomorphism $T^2_{1,k}$ of $C^2_{1,k} \cup \theta(V^2_{1,k})$ onto the same strip. Thus $T_{1,k} = (T^2_{1,k})^{-1}T^1_{1,k}$ on the set $(\lambda(V^1_{1,k}) - \sum \omega^1_{1,k}) \cup \sum \gamma^1_{1,k}$. Further $T^2_{1,k} \cdot T_1 = W$ defines a transformation of $C^1_{1,k}$ onto $y = 0$, $-1 < x < 1$ with the property that

$$(T^1_{1,k})^{-1}W = (T^1_{1,k})^{-1}T^2_{1,k}T_1 = (T_{1,k})^{-1}T_1$$

leaves orientation invariant on $C^1_{1,k}$. It follows, by the lemma proved above, that we can replace $T^1_{1,k}$ by a new o-homeomorphism $\tilde{T}^1_{1,k}$ which has all the properties of $\tilde{T}^1_{1,k}$ plus the property that $\tilde{T}^1_{1,k} = W$ along $C^1_{1,k}$. If we assume $T^1_{1,k}$ chosen as this $\tilde{T}^1_{1,k}$, then along $C^1_{1,k} : T^1_{1,k} = T^2_{1,k}T_1$, or $T_1 = (T^2_{1,k})^{-1}T^{-1}_{1,k} = T_{1,k}$.

Making these restrictions on the $T^1_{1,k}$, we conclude that the transformation T_1 can be extended to a homeomorphism of the set $\lambda(V^1_1) \cup \sum \lambda(V^1_{1,k})$ onto $\lambda(V^2_1) \cup \sum \lambda(V^2_{1,k})$, taking each curve C^1 onto the curve $C^2 = f(C^1)$. Proceeding by induction and making similar restrictions on the T_α, we see that there is an o-homeomorphism T of E^1_0 onto E^2_0 with $T(C^1) = C^2 = f(C^1)$. The same reasoning gives an o-homeomorphism T^* of E^{1*}_0 onto E^{2*}_0 with $T^*(C^1) = C^2 = f(C^1)$. E^{1*}_0 and E^1_0 have the one boundary curve C^1_1 in common. Another application of the above lemma enables us to assume that $T = T^*$ along C^1_1. Hence T can be extended to an o-homeomorphism of the plane onto itself, taking F_1 onto F_2. Theorem 26 is thus established.

We now conclude

THEOREM 27. *If two regular curve-families filling the plane determine isomorphic chordal systems, then they are o-equivalent.*

3. Classification of the curve-families

3.1. **Abstract classification.** Let $\mathfrak{F}$ be the set of all o-equivalence classes $\mathcal{F}$ (see §2.1) of regular curve-families F filling the plane. The relation of isomorphism groups all normal chordal system E in disjoint isomorphism classes $\mathfrak{S}$. Let $\mathfrak{E}$ be the set of all classes $\mathfrak{S}$. On the basis of Theorem I.38 and Theorems 12, 13, 27 we conclude

THEOREM 28. *There is a one-to-one correspondence $w(\mathcal{F}) = \mathfrak{S}$ between the sets $\mathfrak{F}$ and $\mathfrak{E}$ such that, for any F in $\mathcal{F}$ and any E in $\mathfrak{S} = w(\mathcal{F})$, $CS(F)$ is isomorphic to E.*

We have remarked in Part 2, §1 that each full equivalence class $\mathcal{F}^*$ of curve-families generates in general two different o-equivalence classes, $\mathcal{F}$ and $\mathcal{F}'$: $\mathcal{F}^* = \mathcal{F} \cup \mathcal{F}'$. The families of one class, e.g., $\mathcal{F}'$, are obtained by applying one fixed non-orientation-preserving homeomorphism T (for example, $x' = x$, $y' = -y$) to the families of the other class, $\mathcal{F}$. Such a transformation will leave invariant the relations $C_1 \mid C_2 \mid C_3$ but will replace $\mid C_1, C_2, C_3 \mid^+$ by $\mid C'_1, C'_2, C'_3 \mid^-$. In special cases it may happen that the class $\mathcal{F}$ is invariant under such a transformation, in which case $\mathcal{F} = \mathcal{F}' = \mathcal{F}^*$.

We can now introduce a broader equivalence among the normal chordal systems, letting two be *congruent* if they are isomorphic or if one is obtained from the other by a one-to-one transformation which leaves the relation $a \mid b \mid c$ invariant, but replaces $\mid a, b, c \mid^+$ by $\mid a, b, c \mid^-$. We then obtain a one-to-one correspondence between the congruence classes $\mathcal{E}^*$ of normal chordal systems and the equivalence classes $\mathcal{F}^*$ of regular curve-families.

3.2. **Classification by explicit chordal systems.** We first point out that, as a result of Theorem 27 and of Theorem 12 and its proof, we can assert

THEOREM 29. *Let F be a regular curve-family filling the plane. There is an o-homeomorphism of the plane onto the interior of the unit circle such that the family F is transformed onto a regular family F_1 of curves filling the interior of the circle and with the property: each curve of F_1 has a unique pair of limit points on the circumference, no two curves have a common limit point.*

For the family F_1 of curves constructed in the proof of Theorem 12 has the stated property. This family could, moreover, be constructed to have a chordal system isomorphic to any given chordal system, in particular to that of F. By Theorem 27 we can now map F o-homeomorphically on F_1.

If we now join the end-points of each curve of F_1 by a chord, we obtain a family of chords in the circle. No two of these chords can intersect, for then the end-points of one curve of F_1 would separate the end-points of the other curve and the corresponding curves would have to intersect. By Theorem 9 we can immediately introduce the chordal relations in the set of chords in terms of the order of their end-points. This gives a chordal system. It is further isomorphic to CS(F). We have therefore

THEOREM 30. *To each normal chordal system corresponds an isomorphic explicit chordal system.*

We can thus replace the abstract normal chordal systems by the concrete explicit normal chordal systems. It remains to determine when two such explicit chordal systems are isomorphic or congruent:

DEFINITION. Let A and A' be two point sets on the circumference of the unit circle. Let $f(P) = P'$ be a transformation of A on A'. f will be said to be *monotone* if f is one-to-one and if for every ordered triple P_1, P_2, P_3 of points in A the two triangles $P_1P_2P_3$ and $P_1'P_2'P_3'$ have the same orientation.

THEOREM 31. *Let K_1 and K_2 be two explicit chordal systems of the unit circle: $x^2 + y^2 = 1$. Let A_i be the set of end-points of the chords of K_i ($i = 1, 2$). Then K_1 and K_2 are isomorphic if and only if there is a monotone transformation f of A_1 on A_2 such that the image of a pair of end-points of a chord of K_1 is always a pair of end-points of a chord of K_2. K_1 and K_2 are congruent if and only if K_1 is isomorphic either to K_2 or to the image K_2' of K_2 under the transformation $x' = x$, $y' = -y$.*

This theorem follows from the expression of the chordal relations in terms of order of end-points, as in Theorem 9.

Theorem 31 can be regarded as the basis for an *equivalence* (*or o-equivalence*) *criterion* for curve-families. If we imagine an associated isomorphic explicit chordal system K for each family F, then *two families* F_1 *and* F_2 *are equivalent* (*or o-equivalent*) *according as the corresponding systems* K_1 *and* K_2 *satisfy the congruence* (*or isomorphism*) *conditions of Theorem* 31.

APPENDIX. Applications of the classification theory to the cases with singular points

1. The fact that we have made a certain classification of regular curve-families filling the plane, or, what is the same thing, the sphere with one singularity, enables us to attack the case of families on the sphere with many singularities. We give here an illustration.

DEFINITION. An *n-cyclic chordal system* is a chordal system E which is isomorphic to a subset $\lambda(V_1) = V_1 \cup \theta(V_1)$ of a normal chordal system E_1, where V_1 is a cyclic subset of the system E_1.

The class of n-cyclic chordal systems can be studied in detail by means of the general theory of normal chordal systems.

Now take as our example a regular curve-family F filling the sphere with the exception of a closed set A of singular points. Further we assume there is a closed curve Γ on which all points of A lie and that the complementary set $\Gamma - A$ on Γ is composed of curves C_n ($n = 1, 2, \cdots$) of F. The set A need not be countable. We assume it contains at least two points.

To analyze the possible structure of such a family we first consider one region R bounded by Γ. (The theory for the other region is the same.)

First we map Γ plus R homeomorphically on a circle Γ' plus interior R'. Next we join the end-points of each arc C'_n to obtain a chord K'_n. Then (as in the proof of Theorem 12) we transform Γ' plus R' homeomorphically onto the set R'' bounded by the curve Γ'' formed by the chords K'_n and the set A'. The subset F_0 of F in $\Gamma \cup R$ is thus mapped homeomorphically on a curve-family F'_0 filling the interior of the curve Γ'' plus the chords K'_n. We extend F'_0 to a family filling the interior of the circle simply by filling out the segments bounded by the K'_n by chords parallel to K'_n. The result is a regular family filling the interior of the circle.

We conclude that the curve-family F'_0 is n-cyclic. But F'_0 is a homeomorphic image of F_0. Hence the structure of F_0 is the same as that of an n-cyclic family, to which our previous theory can be applied.

The following converse can be established, by means of recent results of Adkisson and MacLane (see this Journal, vol. 6(1940), p. 216 ff.): given any n-cyclic chordal system $E = V + \theta(V)$, whose "boundary" V is isomorphic to the set V' of chords K'_n, then there is a regular curve-family F filling R plus the C_n whose structure is that of E; i.e., F'_0 is isomorphic to E.

2. Generalization. Suppose now that we have a regular curve-family F filling the sphere except for a closed point set A which lies on a closed set Γ such that each component region R_i of the complement of Γ on the sphere is a simply-connected region bounded by a closed curve $\Gamma_i \subset \Gamma$ and that $\Gamma - A$ is formed of curves of F. To each region $R_i \cup \Gamma_i$ can then be applied the same reasoning as above. We obtain again n-cyclic families.

UNIVERSITY OF MICHIGAN.

THE FRACTIONAL DERIVATIVE OF A LAPLACE INTEGRAL

By C. V. L. Smith

Introduction. The integral

$$\int_0^\infty e^{-xt}\, d\alpha(t), \tag{1}$$

where $\alpha(t)$, a function of the real variable t, is of bounded variation in $(0, R)$ for every positive R, has been exhaustively studied by D. V. Widder [9, 10, 11, 12].[1] In the present paper, we shall consider the integral

$$\int_0^\infty e^{-xt}\, t^\rho\, d\alpha(t), \tag{2}$$

where $\alpha(t)$ is as described above, and ρ is a positive constant, restricting ourselves to the case where x and $\alpha(t)$ are real.

In the case of the integral (1), $\alpha(t)$ is said to be normalized if

$$\alpha(0) = 0, \qquad \alpha(t) = \frac{\alpha(t+) + \alpha(t-)}{2} \qquad (0 < t < \infty). \tag{3}$$

In the present case, we may also take $\alpha(0+) = 0$. For let $\beta(0) = 0$, $\beta(t) = \alpha(0+)$ for $0 < t$, and set $\alpha^*(t) = \alpha(t) - \beta(t)$ for $0 \leqq t$. Since the integral

$$\int_0^\infty e^{-xt}\, t^\rho\, d\beta(t) \tag{4}$$

is obviously convergent for all real x and has the value zero, it is clear that wherever the integral (2) converges we have

$$\int_0^\infty e^{-xt}\, t^\rho\, d\alpha(t) = \int_0^\infty e^{-xt}\, t^\rho\, d\alpha^*(t). \tag{5}$$

The function $\alpha(t)$ will always be taken as satisfying the conditions (3) and the condition $\alpha(0+) = 0$; and any function satisfying these conditions will be said to be normalized.

The derivatives of the function defined by a convergent integral of the form (1) are given by the integrals ([9], p. 702)

$$(-1)^k \int_0^\infty e^{-xt}\, t^k\, d\alpha(t) \qquad (k = 1, 2, \cdots), \tag{6}$$

which leads one to expect that, if ρ is non-integral, (2) is either the fractional derivative of order ρ of (1) or its negative. Integrals of the form (1) con-

Received May 15, 1940.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

verge for $c < x < \infty$, c being a constant ([9], p. 700), so that neither the Riemann nor the Weyl definitions[2] of the fractional derivative are applicable in the present case. However, it is possible to define fractional differentiation in such a way that (2) is $(-1)^{-m}$ times the ρ-th derivative of (1), where m is the greatest integer which does not exceed ρ. We first define fractional integration by the formula

$$(7) \qquad {}_{\infty}I_x^{\rho}[f(x)] = \frac{1}{\Gamma(\rho)} \int_x^{\infty} (u - x)^{\rho-1} f(u)\, du$$

in case the integral converges. Now let m be the greatest integer which does not exceed ρ, and set $\nu = \rho - m$. We define the ρ-th order derivative of $f(x)$ as the negative of the $(1 - \nu)$-th integral of the $(m + 1)$-th derivative, that is, as follows:

$$(8) \qquad {}_{\infty}D_x^{\rho}[f(x)] = \frac{-1}{\Gamma(1 - \nu)} \int_x^{\infty} (u - x)^{-\nu} f^{(m+1)}(u)\, du.$$

This definition is justified in the case of functions defined by integrals of the form (1), for here ([11], p. 119) $f^{(k)}(x) = o(x^{-k})$ as x becomes infinite for $k = 1, 2, \cdots$, and hence, if ρ is a positive integer, we have $\nu = 0$, and

$$(9) \qquad {}_{\infty}D_x^{\rho}[f(x)] = -\int_x^{\infty} f^{(\rho+1)}(u)\, du = f^{(\rho)}(x).$$

It will be shown that if $g(x)$ is the function defined by a convergent integral of the form (1), and $f(x)$ is the function defined by the corresponding integral of the form (2), we have

$$(10) \qquad f(x) = (-1)^{-m}\, {}_{\infty}D_x^{\rho}[g(x)]$$

and

$$(11) \qquad g(x) = {}_{\infty}I_x^{\rho}[f(x)].$$

For negative values of ρ, the convergence of (1) does not imply that of (2), as we may see by taking

$$(12) \qquad \alpha(0) = 0, \qquad \alpha(t) = \int_0^t \frac{du}{u[\log u/(1 + u)]^2}, \qquad 0 < t,$$

and so we obviously cannot obtain (2) by a fractional operation on (1). It is for this reason that we have omitted a consideration of the integral (2) in the case where ρ is negative.

In the second part of the paper, a means of inverting the integral (2) is found. To do this, we employ the operator ([11], p. 117)

$$(13) \qquad L_{k,t}(f(x)) = \frac{(-1)^k}{k!} \left(\frac{k}{t}\right)^{k+1} f^{(k)}\left(\frac{k}{t}\right) \qquad (0 < t < \infty;\ k = 1, 2, \cdots),$$

[2] For the Riemann and the Weyl definitions of fractional integration, upon which the definitions of fractional differentiation depend, see, for example, [13], pp. 222, 224.

by means of which the inversion of (1) has been effected. In the present case, the inversion formula is found to be

$$\alpha(t) = \lim_{k\to\infty} \int_0^t u^{-\rho} L_{k,u}(f(x))\, du \qquad (0 < t < \infty), \tag{14}$$

if $\alpha(t)$ is normalized in the sense we are using.

In the third part of the paper, we shall find sets of necessary and sufficient conditions that a function may have an integral representation of the form (2) with $\alpha(t)$ belonging to various functional classes.

Part I

1. **General properties.** An integral of the form (2) is said to converge for a value of x if for that value

$$\lim_{R\to\infty} \int_0^R e^{-xt} t^\rho\, d\alpha(t) \tag{1.1}$$

exists. By an important theorem on the Stieltjes integral ([3], p. 11), we may write

$$\int_0^R e^{-xt} t^\rho\, d\alpha(t) = \int_0^R e^{-xt}\, d\beta(t) \qquad (0 < R < \infty), \tag{1.2}$$

where

$$\beta(0) = 0, \qquad \beta(t) = \int_0^t u^\rho\, d\alpha(u) \qquad (0 < t < \infty). \tag{1.3}$$

Hence

$$\int_0^\infty e^{-xt} t^\rho\, d\alpha(t) = \int_0^\infty e^{-xt}\, d\beta(t) \tag{1.4}$$

for all values of x for which the limit (1.1) exists. The integral

$$\int_0^\infty e^{-xt}\, d\beta(t) \tag{1.5}$$

being of the form (1), either converges for all values of x, or diverges for all values of x, or there exists a number c such that (1.5) converges for $c < x$ and diverges for $x < c$, the number c being said to be the abscissa of convergence of (1.5) ([9], p. 700). By (1.4), it is clear that the same statement can be made about the convergence of (2).

Theorem 1.1. *Let ρ be a positive constant, and let $\alpha(t)$ be a function of bounded variation in* $(0, R)$ *for every positive R, then the integrals* (1) *and* (2) *either both converge for all values of x, or both diverge for all values of x, or have the same abscissa of convergence.*

Suppose (1) converges for $x > c$. Given $\epsilon > 0$, we have

$$(1.6) \qquad \lim_{\substack{R,S\to\infty \\ S\geqq R}} \int_R^S e^{-(c+\epsilon)t}\, d\alpha(t) = 0.$$

For $t > \rho/\epsilon$, the function $t^\rho e^{-\epsilon t}$ is bounded and monotonic decreasing, and so, for $R > \rho/\epsilon$, we have the inequality

$$(1.7) \qquad \left| \int_R^S e^{-(c+\epsilon)t} t^\rho e^{-\epsilon t}\, d\alpha(t) \right| \leqq R^\rho e^{-\epsilon R} \sup_{R\leqq T\leqq S} \left| \int_R^T e^{-(c+\epsilon)t}\, d\alpha(t) \right|$$

by the second law of the mean for Stieltjes integrals.[3] By (1.6) and the fact that $R^\rho e^{-\epsilon R}$ is monotonic decreasing, it follows that the right side of (1.7) approaches zero as R and S become infinite. Hence (2) converges for $x = c + 2\epsilon$, where $\epsilon > 0$ is arbitrary, and so for $x > c$.

On the other hand, let (2) converge for $x > c$. For $\epsilon > 0$ we have

$$(1.8) \qquad \lim_{\substack{R,S\to\infty \\ S\geqq R}} \int_R^S e^{-(c+\epsilon)t} t^\rho\, d\alpha(t) = 0.$$

But for $R > 0$,

$$(1.9) \qquad \begin{aligned} \left| \int_R^S e^{-(c+\epsilon)t}\, d\alpha(t) \right| &= \left| \int_R^S e^{-(c+\epsilon)t} t^{-\rho} t^\rho\, d\alpha(t) \right| \\ &\leqq R^{-\rho} \sup_{R\leqq T\leqq S} \left| \int_R^T e^{-(c+\epsilon)t} t^\rho\, d\alpha(t) \right| \end{aligned}$$

by an application of the second law of the mean for Stieltjes integrals. By (1.8), it follows that (1) converges for $x = c + \epsilon$, and so for $x > c$.

If (1) converges for no value of x, then clearly (2) converges for no value of x. For if (2) were to converge for $x > c$, (1) would converge there also. Similarly, if (2) converges for no value of x, the same is true of (1).

If (1) converges for all x, let x_0 be an arbitrary fixed value of x. It follows from what we have just proved that (2) converges for $x > x_0$, and hence for all values of x, since x_0 is arbitrary. Similarly, if (2) converges for all values of x, so does (1).

Finally, suppose (1) has the abscissa of convergence c. Then by our argument above (2) converges for $x > c$, so that the abscissa of convergence of (2) is c_1 where $c_1 \leqq c$. But we cannot have $c_1 < c$, for in that case (1) would converge for $c_1 < x < c$, and this is contrary to hypothesis. Hence we must have $c_1 = c$. Similarly, if c is the abscissa of convergence of (2), then c is also the abscissa of convergence of (1).

COROLLARY 1.11. *The abscissa of convergence of* (2), *if it is positive, is given by* ([9], p. 704)

$$(1.10) \qquad c = \limsup_{t\to\infty} \frac{\log |\alpha(t)|}{t}.$$

[3] We are using the lemma on page 6 of [1] with a slight change in notation.

COROLLARY 1.12. *If a constant γ exists such that*

$$|\alpha(t)| < e^{\gamma t} \qquad (0 \leqq t < \infty), \tag{1.11}$$

then (2) *converges for* $\gamma < x$. (See [9], p. 703.)

COROLLARY 1.13. *If* (2) *converges for* $x = x_0 > 0$, *then a constant* $K > 0$ *exists such that*

$$|\alpha(t)| < Ke^{x_0 t} \qquad (0 \leqq t < \infty). \tag{1.12}$$

(See [9], p. 703.)

2. The fractional derivative and integral.

DEFINITION 2.1. Let ρ be a positive number. Then the letter m is to designate the greatest positive integer which does not exceed ρ, and the letter ν is to designate the difference $\rho - m$.

DEFINITION 2.2. The fractional integral of order $\rho > 0$ of a function $f(x)$ is defined by the equation

$$_\infty I_x^\rho[f(x)] = \frac{1}{\Gamma(\rho)} \int_x^\infty (u - x)^{\rho - 1} f(u)\, du \tag{2.1}$$

provided the integral converges.

DEFINITION 2.3. The fractional derivative of order $\rho > 0$ of a function $f(x)$ is defined by the equation

$$_\infty D_x^\rho(f(x)) = -\frac{1}{\Gamma(1 - \nu)} \int_x^\infty (u - x)^{-\nu} f^{(m+1)}(u)\, du \tag{2.2}$$

provided the integral converges. Here m and ν are used in the sense of Definition 2.1.

THEOREM 2.1. *Let $\alpha(t)$ be a function of bounded variation in $(0, R)$ for every positive R, and such that the abscissa of convergence of the integrals* (1) *and* (2) *is c. Let $g(x)$ be the function defined by the integral* (1), *and let $f(x)$ be the function defined by the integral* (2). *Then $f(x)$ is given by the equation*

$$f(x) = (-1)^{-m}\, _\infty D_x^\rho(g(x)) \qquad (c < x < \infty). \tag{2.3}$$

It must first be shown that the fractional derivative exists. We have ([9], p. 702)

$$g^{(m+1)}(u) = (-1)^{m+1} \int_0^\infty e^{-ut} t^{m+1}\, d\alpha(t) \qquad (c < u < \infty), \tag{2.4}$$

where $g^{(m+1)}(u) = o(u^{-m-1})$ as u becomes infinite ([11], p. 119). If $0 < \rho < 1$, we have $m = 0$ and $\nu = \rho$. It is clear, then, that the integral

$$\int_x^\infty (u - x)^{-\nu} g^{(m+1)}(u)\, du \qquad (c < x < \infty) \tag{2.5}$$

is absolutely convergent, so that the fractional derivative of order ρ of $g(x)$ exists for $c < x < \infty$.

We first consider the case where $c \geqq 0$. By an application of Stieltjes integration by parts ([6], p. 539) we obtain the following equation:

$$\int_0^R e^{-ut} t^{m+1} \, d\alpha(t) = e^{-uR} R^{m+1} \alpha(R) + \int_0^R e^{-ut} (ut^{m+1} - (m+1)t^m) \alpha(t) \, dt \tag{2.6}$$

$$(c < u < \infty, 0 < R < \infty).$$

Making use of Corollary 1.13, we obtain, upon allowing R to become infinite in (2.6),

$$\int_0^\infty e^{-ut} t^{m+1} \, d\alpha(t) = \int_0^\infty e^{-ut} [ut^{m+1} - (m+1)t^m] \alpha(t) \, dt \quad (c < u < \infty). \tag{2.7}$$

Setting $u = x + y$ in (2.5) and making use of (2.7), we obtain the following equation for ${}_\infty D_x^\rho(g(x))$:

$$\begin{aligned} {}_\infty D_x^\rho(g(x)) = \frac{(-1)^m}{\Gamma(1-\nu)} \Bigg[& \int_0^\infty y^{1-\nu} \, dy \int_0^\infty e^{-(y+x)t} t^{m+1} \alpha(t) \, dt \\ & + x \int_0^\infty y^{-\nu} \, dy \int_0^\infty e^{-(y+x)t} t^{m+1} \alpha(t) \, dt \\ & - (m+1) \int_0^\infty y^{-\nu} \, dy \int_0^\infty e^{-(y+x)t} t^m \alpha(t) \, dt \Bigg] \quad (c < x < \infty). \end{aligned} \tag{2.8}$$

Now let $x_0 > c$ be an arbitrary, but fixed, value of x, and let x_1 be such that $c < x_1 < x_0$. By Corollary 1.13 it is clear that for $x = x_0$ each of the iterated integrals on the right of equation (2.8) remains convergent if its integrand is replaced by its absolute value. Therefore, by an application of the Fubini theorem ([7], p. 347), we have the following equation

$$\begin{aligned} {}_\infty D_x^\rho(g(x)) \,|_{x=x_0} = \frac{(-1)^m}{\Gamma(1-\nu)} \Bigg[& \int_0^\infty e^{-x_0 t} t^{m+1} \alpha(t) \, dt \int_0^\infty e^{-yt} y^{1-\nu} \, dy \\ & + x_0 \int_0^\infty e^{-x_0 t} t^{m+1} \alpha(t) \, dt \int_0^\infty e^{-yt} y^{-\nu} \, dy \\ & - (m+1) \int_0^\infty e^{-x_0 t} t^m \alpha(t) \, dt \int_0^\infty e^{-yt} y^{-\nu} \, dy \Bigg] \end{aligned} \tag{2.9}$$

which becomes, when the integrations with respect to y have been performed,

$$\begin{aligned} {}_\infty D_x^\rho(g(x)) \,|_{x=x_0} = \frac{(-1)^m}{\Gamma(1-\nu)} \Bigg[& \Gamma(2-\nu) \int_0^\infty e^{-x_0 t} t^{\rho-1} \alpha(t) \, dt \\ & + x_0 \Gamma(1-\nu) \int_0^\infty e^{-x_0 t} t^\rho \alpha(t) \, dt - (m+1)\Gamma(1-\nu) \int_0^\infty e^{-x_0 t} t^{\rho-1} \alpha(t) \, dt \Bigg]. \end{aligned} \tag{2.10}$$

But $\Gamma(2 - \nu) = (1 - \nu)\Gamma(1 - \nu)$, so that, if we perform the division by $\Gamma(1 - \nu)$ in the right member of (2.10) and collect terms, we obtain

$$
\begin{aligned}
{}_\infty D_x^\rho(g(x))\,|_{x=x_0} &= (-1)^m \left[x_0 \int_0^\infty e^{-x_0 t} t^\rho \alpha(t)\,dt - \rho \int_0^\infty e^{-x_0 t} t^{\rho-1} \alpha(t)\,dt \right] \\
&= (-1)^m \int_0^\infty e^{-x_0 t} t^\rho \,d\alpha(t).
\end{aligned} \tag{2.11}
$$

Since x_0 is an arbitrary number greater than c, the truth of the theorem follows from (2.11) in the case where c is non-negative.

Now suppose c is negative. Here we cannot pass from (2.6) to (2.7) as in the previous case. However, we have

$$
g(x) = \int_0^\infty e^{-xt}\,d\alpha(t) = \int_0^\infty e^{-(x-c)t} e^{-ct}\,d\alpha(t) = \int_0^\infty e^{-(x-c)t}\,d\beta(t) \qquad (c < x < \infty), \tag{2.12}
$$

where

$$
\beta(0) = 0, \qquad \beta(t) = \int_0^t e^{-cu}\,d\alpha(u) \qquad (0 < t < \infty), \tag{2.13}
$$

so that $\beta(t)$ is of bounded variation in $(0, R)$ for every positive R and is such that

$$
\int_0^\infty e^{-(x-c)t}\,d\beta(t) \tag{2.14}
$$

converges for $x - c > 0$. We set $y = x - c$ in (2.14), and designate the function defined by the resulting integral by $G(y)$. In a similar fashion we obtain $F(y)$ from $f(x)$. Since the abscissa of convergence of the integrals defining $G(y)$ and $F(y)$ is zero, we have, by the first part of the proof,

$$
(-1)^{-m}\,{}_\infty D_y^\rho[G(y)] = F(y) \qquad (0 < y < \infty). \tag{2.15}
$$

We have $F(y) = f(x)$, where $y = x - c$, since

$$
\int_0^\infty e^{-(x-c)t} t^\rho\,d\beta(t) = \int_0^\infty e^{-xt} t^\rho e^{ct}\,d\beta(t) = \int_0^\infty e^{-xt} t^\rho\,d\alpha(t) \qquad (0 < y = x - c < \infty), \tag{2.16}
$$

and similarly $G^{(m+1)}(y) = g^{(m+1)}(x)$ $(0 < y = x - c < \infty)$. Furthermore, we have

$$
\begin{aligned}
\int_y^\infty (u - y)^{-\nu} G^{(m+1)}(u)\,du &= \int_{x-c}^\infty (u - x + c)^{-\nu} G^{(m+1)}(u)\,du \\
&= \int_x^\infty (u - x)^{-\nu} g^{(m+1)}(u)\,du \quad (0 < y = x - c < \infty).
\end{aligned} \tag{2.17}
$$

From (2.15), (2.16), and (2.17), we obtain (2.3) in this case also. This completes the proof.

THEOREM 2.2. *Let $\alpha(t)$ be a normalized function of bounded variation in* (0, R) *for every positive R, and such that the integrals* (1) *and* (2) *have the abscissa of convergence c. Let $f(x)$ be the function defined by the integral* (2), *and let $g(x)$ be the function defined by the integral* (1). *Then $g(x)$ is given by the formula*

$$g(x) = {}_{\infty}I_x^{\rho}(f(x)) \qquad (c < x < \infty). \tag{2.18}$$

We first consider the case $c \geqq 0$. The integrals

$$\int_0^R |L_{k,t}(g(x))| \, dt \qquad (0 < R < \infty; k = 1, 2, \cdots) \tag{2.19}$$

all exist (for $L_{k,t}(g(x))$ see the introduction, (13); see [12], p. 251). In (2.19), set $k = m + 1$ and $R = (m + 1)/x$ where $c < x$, and introduce the variable u by setting $u = (m + 1)/t$. Thus we are assured of the convergence of the integral

$$\int_x^\infty \frac{u^m}{m!} | g^{(m+1)}(u) | \, du \qquad (c < x < \infty). \tag{2.20}$$

Now consider the integral

$$\int_x^u (u - y)^{-\nu}(y - x)^{\rho-1} \, dy \qquad (c < x \leqq u < \infty). \tag{2.21}$$

In (2.21), introduce the variable λ by the equation

$$y = \lambda(u - x) + x \tag{2.22}$$

so that (2.21) becomes

$$(u - x)^m \int_0^1 \lambda^{\rho-1}(1 - \lambda)^{-\nu} \, d\lambda = (u - x)^m B(\rho, 1 - \nu) \qquad (c < x \leqq u < \infty). \tag{2.23}$$

The integral

$$\int_x^\infty \frac{(u - x)^m}{m!} | g^{(m+1)}(u) | \, du \qquad (c < x < \infty) \tag{2.24}$$

converges, as we may see ([2], p. 429) by multiplying the integrand of the convergent integral (2.20) by the bounded monotonic factor $[(u - x)/u]^m$. Hence the integral

$$\int_x^\infty \frac{| g^{(m+1)}(u) |}{m!} \int_x^\infty G(u, y)(u - y)^{-\nu}(y - x)^{\rho-1} \, dy \quad (c < x < \infty), \tag{2.25}$$

where

$$G(u, y) = \begin{cases} 0 & (u < y), \\ 1 & (u \geqq y) \end{cases} \tag{2.26}$$

is convergent. Therefore, by the Fubini theorem, we have

$$\frac{(-1)^m B(\rho, 1-\nu)}{m!}\int_x^\infty (u-x)^m g^{(m+1)}(u)\,du$$

$$(2.27)\qquad = (-1)^m \int_x^\infty \frac{g^{(m+1)}(u)\,du}{m!}\int_x^u G(u,y)(u-y)^{-\nu}(y-x)^{\rho-1}\,dy$$

$$= \frac{(-1)^m}{m!}\int_x^\infty (y-x)^{\rho-1}\,dy \int_y^\infty (u-y)^{-\nu} g^{(m+1)}(u)\,du$$

$$(c < x < \infty),$$

so that, since

$$(2.28)\qquad B(\rho, 1-\nu) = \frac{\Gamma(\rho)\Gamma(1-\nu)}{\Gamma(\rho+1-\nu)} = \frac{\Gamma(\rho)\Gamma(1-\nu)}{m!},$$

we finally obtain the equation

$$\frac{(-1)^m}{m!}\int_x^\infty (u-x)^m g^{(m+1)}(u)\,du$$

$$(2.29)\qquad = \frac{(-1)^m}{\Gamma(\rho)}\int_x^\infty (y-x)^{\rho-1}\,dy\int_y^\infty \frac{(u-y)^{-\nu}}{\Gamma(1-\nu)}\, g^{(m+1)}(u)\,du.$$

Since $g^{(n)}(u) = o(u^{-n})$ as u becomes infinite for $n = 1, 2, \cdots$ ([11], p. 119), it is clear that the integral on the left of (2.29) can be integrated by parts m times in succession. Therefore, since $g(\infty) = \alpha(0+)$ ([11], p. 120) and we are taking $\alpha(0+) = 0$, we have

$$(2.30)\qquad \frac{(-1)^m}{m!}\int_x^\infty (u-x)^m g^{(m+1)}(u)\,du = \int_x^\infty g'(u)\,du = -g(x) \qquad (c < x < \infty).$$

From (2.30), (2.29), the definitions of fractional integration and differentiation, and Theorem 2.1, we obtain (2.18) in the case $c \geqq 0$. The theorem can be established for the case $c < 0$ by the method used in the proof of Theorem 2.1 for this case; the details need not be given here.

3. The derivatives of $f(x)$.

THEOREM 3.1. *The derivatives of a function $f(x)$ defined by a convergent integral of the form* (2) *are given by the equation*

$$(3.1)\qquad f^{(k)}(x) = (-1)^k\int_0^\infty e^{-xt}t^{k+\rho}\,d\alpha(t) \qquad (c < x < \infty;\, k = 1, 2, \cdots).$$

The proof, which the reader will have no difficulty in supplying, will be omitted.

COROLLARY 3.11. *Under the conditions of Theorem* 2.1, *the derivatives of $f(x)$ are given by the formula*

$$(3.2)\qquad f^{(k)}(x) = (-1)^{-m}\,{}_\infty D_x^\rho(g^{(k)}(x)) \qquad (c < x < \infty;\, k = 1, 2, \cdots).$$

The proof, which follows exactly the same method as that of Theorem 2.1, need not be given here.

COROLLARY 3.12. *Under the conditions of Theorem 2.2, the derivatives of $g(x)$ are given by the formula*

$$g^{(k)}(x) = {}_\infty I_x^\rho[f^{(k)}(x)] \qquad (c < x < \infty;\ k = 1, 2, \cdots). \tag{3.3}$$

The proof follows the same method as that of Theorem 2.2, and will not be given here.

4. The order of $f^{(k)}(x)$ as x becomes infinite.

THEOREM 4.1. *If $f(x)$ is the function defined by a convergent integral of the form (2), then*

$$f^{(k)}(x) = o(x^{-k-\rho}) \qquad (x \to \infty;\ k = 0, 1, 2, \cdots). \tag{4.1}$$

By Theorem 2.1 and Corollary 3.11, we have

$$|f^{(k)}(x)| \leqq \frac{1}{\Gamma(1-\nu)} \int_x^\infty (u-x)^{-\nu} |g^{(k+m+1)}(u)|\, du \qquad (c < x < \infty;\ k = 0, 1, 2, \cdots). \tag{4.2}$$

Since $g^{(n)}(u) = o(u^{-n})$ as u becomes infinite for $n = 1, 2, \cdots$, we can, for any particular value of k, given an $\epsilon > 0$, determine x_0 such that

$$|f^{(k)}(x)| < \frac{\epsilon}{\Gamma(1-\nu)} \int_x^\infty (u-x)^{-\nu} u^{-k-m-1}\, du \qquad (x_0 < x < \infty). \tag{4.3}$$

By the change of variable $u = x/\lambda$, we obtain

$$\begin{aligned} \int_x^\infty (u-x)^{-\nu} u^{-k-m-1}\, du &= x^{-\rho-k} \int_0^1 (1-\lambda)^{-\nu} \lambda^{k+\rho-1}\, d\lambda \\ &= x^{-\rho-k} B(1-\nu, k+\rho). \end{aligned} \tag{4.4}$$

From (4.3) and (4.4), (4.1) immediately follows.

5. The uniqueness of $\alpha(t)$.

THEOREM 5.1. *If $f(x)$ has two representations of the form (2), where ρ is the same positive number in both cases,*

$$\begin{aligned} &\text{(a)} \quad f(x) = \int_0^\infty e^{-xt} t^\rho\, d\alpha_1(t) \qquad (c < x < \infty), \\ &\text{(b)} \quad f(x) = \int_0^\infty e^{-xt} t^\rho\, d\alpha_2(t) \qquad (c < x < \infty), \end{aligned} \tag{5.1}$$

where $\alpha_1(t)$ and $\alpha_2(t)$ are normalized functions of bounded variation in $(0, R)$ for every positive R, then

$$\alpha_1(t) = \alpha_2(t) \qquad (0 \leqq t < \infty). \tag{5.2}$$

We recall that a normalized function, in the sense in which we are using the term, satisfies conditions (3) of the introduction, and the condition $\alpha(0+) = 0$. Hence, if we set $\gamma(t) = \alpha_1(t) - \alpha_2(t)$, then $\gamma(t)$ is also a normalized function, and we clearly have

$$\int_0^\infty e^{-xt} t^\rho \, d\gamma(t) = 0 \qquad (c < x < \infty). \tag{5.3}$$

From this it follows, if we set

$$\beta(0) = 0, \qquad \beta(t) = \int_1^t u^\rho \, d\gamma(u) \qquad (0 < t < \infty), \tag{5.4}$$

that

$$\int_0^\infty e^{-xt} \, d\beta(t) = 0 \qquad (c < x < \infty). \tag{5.5}$$

Then $\beta(t)$ vanishes except on a set of measure zero ([9], p. 706), and therefore, since clearly $\beta(t)$ is a normalized function, it follows that $\beta(t)$ vanishes identically. But we have

$$\int_1^t d\gamma(u) = \int_1^t u^{-\rho} \, d\beta(u) = 0 \qquad (0 < t < \infty), \tag{5.6}$$

so that $\gamma(t) = \gamma(1)$, and hence, since we must have $\gamma(1) = \gamma(0+) = 0$, the truth of (5.2) follows.

Part II

6. **The inversion formula when $\alpha(t)$ is absolutely continuous.** If $\alpha(t)$ is absolutely continuous in $(0, R)$ for every positive R, then

$$\alpha(t) = \int_0^t \phi(u) \, du \qquad (0 < t < \infty), \tag{6.1}$$

where $\phi(t)$ is integrable in the sense of Lebesgue in $(0, R)$ for every positive R. Then we may write ([6], p. 665)

$$\int_0^R e^{-xt} t^\rho \, d\alpha(t) = \int_0^R e^{-xt} t^\rho \phi(t) \, dt \qquad (0 < R < \infty). \tag{6.2}$$

Provided $c < x$, the integral on the left of (6.2) approaches a limit as R becomes infinite. Hence the same is true of the integral on the right of (6.2), and the following equation holds:

$$\int_0^\infty e^{-xt} t^\rho \, d\alpha(t) = \int_0^\infty e^{-xt} t^\rho \phi(t) \, dt \qquad (c < x < \infty). \tag{6.3}$$

Conversely, if $\phi(t)$ is integrable in the sense of Lebesgue in $(0, R)$ for every positive R, and if

$$\int_0^\infty e^{-xt} t^\rho \phi(t) \, dt \tag{6.4}$$

exists, even as a Cauchy-value, for $c < x$, then we may define $\alpha(t)$ by equation (6.1), and write (6.2), from which (6.3) follows because in this case the integral on the right of (6.2) approaches a limit as R becomes infinite.

We first obtain an inversion formula for integrals of the form (6.4), and then proceed to the case of the general integral of the form (2). Both inversion formulas involve the operator $L_{k,t}(f(x))$ which has been defined in (13) of the introduction.

THEOREM 6.1. *If $f(x)$ is the function defined by a convergent integral of the form* (6.4), *where $\phi(t)$ is of class L in* (0, *R*) *for every positive R, then*

$$\lim_{k\to\infty} t^{-\rho} L_{k,t}(f(x)) = \phi(t) \tag{6.5}$$

at all points of the Lebesgue set (cf. [8], p. 364) *of $\phi(t)$.*

If we set $t^\rho\phi(t) = \psi(t)$, then we have

$$f(x) = \int_0^\infty e^{-xt}\psi(t)\,dt, \tag{6.6}$$

where $\psi(t)$ is of class L in $(0, R)$ for every positive R. In this case the inversion formula is ([11], p. 122)

$$\lim_{k\to\infty} L_{k,t}(f(x)) = \psi(t) = t^\rho\phi(t) \tag{6.7}$$

at all points of the Lebesgue set of $\psi(t)$, and hence evidently at all points of the Lebesgue set of $\phi(t)$. The theorem clearly follows from (6.7).

7. The general case.

THEOREM 7.1. *If $f(x)$ is the function defined by a convergent integral of the form* (2), *where $\alpha(t)$ is normalized, then*

$$\alpha(t) = \lim_{k\to\infty}\int_0^t u^{-\rho}L_{k,u}(f(x))\,du \qquad (0 < t < \infty). \tag{7.1}$$

Let t be an arbitrary positive number, which will be fixed in the course of the argument. Let c be the abscissa of convergence of (2), and let c_1 be the greater of the two numbers c and zero. Then $L_{k,u}(f(x))$ will certainly exist for $0 < u \leqq t$ and $k > c_1t$. The expression $u^{-\rho}L_{k,u}(f(x))$ has the value

$$u^{-\rho}L_{k,u}(f(x)) = \frac{u^{-\rho}}{k!}\left(\frac{k}{u}\right)^{k+1}\int_0^\infty e^{-ky/u}y^{k+\rho}\,d\alpha(y) \quad (0 < u \leqq t;\ k > c_1t), \tag{7.2}$$

where we have substituted for $f^{(k)}(k/u)$ its value as given by Theorem 3.1. The right member of equation (7.2) may be written in the form

$$u^{-1}\frac{k^{k+1}}{k!}\int_0^\infty e^{-ky/u}\left[\frac{y^{k+\rho}}{u^{k+\rho}}\right]d\alpha(y) \quad (0 < u < t;\ k > c_1t). \tag{7.3}$$

If Stieltjes integration by parts is applied to the integral in the expression (7.3), we obtain

$$-u^{-1}\frac{k^{k+1}}{k!}\int_0^\infty \alpha(y)\frac{\partial}{\partial y}\left[e^{-ky/u}\frac{y^{k+\rho}}{u^{k+\rho}}\right]dy \qquad (0 < u < t;\, k > c_1 t). \tag{7.4}$$

The function

$$e^{-ky/u}\frac{y^{k+\rho}}{u^{k+\rho}} \tag{7.5}$$

is homogeneous of degree zero, and possesses continuous first partial derivatives for $0 \leqq y < \infty$, $0 < u < \infty$. Therefore, by Euler's theorem, the following equation holds:

$$y\frac{\partial}{\partial y}\left[e^{-ky/u}\frac{y^{k+\rho}}{u^{k+\rho}}\right] + u\frac{\partial}{\partial u}\left[e^{-ky/u}\frac{y^{k+\rho}}{u^{k+\rho}}\right] = 0 \quad (0 \leqq y < \infty;\, 0 < u < \infty). \tag{7.6}$$

In consequence of (7.6), (7.4) becomes

$$\frac{k^{k+1}}{k!}\int_0^\infty \alpha(y)y^{k+\rho-1}\frac{\partial}{\partial u}\left[\frac{e^{-ky/u}}{u^{k+\rho}}\right]dy \quad (0 < u < t;\, k > c_1 t). \tag{7.7}$$

Now let δ be a positive number such that $0 < \delta < \frac{1}{2}t$. Since

$$\frac{\partial}{\partial u}\left[\frac{e^{-ky/u}}{u^{k+\rho}}\right] = \frac{[ky - (k+\rho)u]e^{-ky/u}}{u^{k+\rho+2}} \qquad (0 \leqq y < \infty;\, 0 < u < \infty), \tag{7.8}$$

it is easy to see that

$$\frac{k^{k+1}}{k!}\int_\delta^t du\int_0^\infty \left|\alpha(y)y^{k+\rho-1}\frac{\partial}{\partial u}\left[\frac{e^{-ky/u}}{u^{k+\rho}}\right]\right| dy \qquad (0 < \delta < \tfrac{1}{2}t;\, k > c_1 t) \tag{7.9}$$

converges. Hence, by the Fubini theorem, we can change the order of integration in the convergent integral

$$\frac{k^{k+1}}{k!}\int_\delta^t du\int_0^\infty \alpha(y)y^{k+\rho-1}\frac{\partial}{\partial u}\left[\frac{e^{-ky/u}}{u^{k+\rho}}\right]dy \qquad (0 < \delta < \tfrac{1}{2}t;\, k > c_1 t) \tag{7.10}$$

and obtain, making use of (7.2), the equation

$$\int_\delta^t u^{-\rho}L_{k,u}(f(x))\,du = \frac{k^{k+1}}{k!}\int_0^\infty \alpha(y)y^{k+\rho-1}\,dy\int_\delta^t \frac{\partial}{\partial u}\left[\frac{e^{-ky/u}}{u^{k+\rho}}\right]du \qquad (0 < \delta < \tfrac{1}{2}t;\, k > c_1 t). \tag{7.11}$$

If the integration with respect to u in (7.11) is carried out, we get

$$\begin{aligned}\int_\delta^t u^{-\rho}L_{k,u}(f(x))\,du = {} & \frac{k^{k+1}}{k!}\int_0^\infty \frac{e^{-ky/t}}{t^{k+\rho}}\,y^{k+\rho-1}\alpha(y)\,dy \\ & - \frac{k^{k+1}}{k!}\int_0^\infty \frac{e^{-ky/\delta}}{\delta^{k+\rho}}\,y^{k+\rho-1}\alpha(y)\,dy \\ & (0 < \delta < \tfrac{1}{2}t;\, k > c_1 t).\end{aligned} \tag{7.12}$$

We shall eventually allow δ to approach zero in (7.12). It will first be shown that

$$\int_0^t |u^{-\rho} L_{k,u}(f(x))| \, du \qquad (k > c_1 t) \tag{7.13}$$

exists.

Consider the function $g(x)$ of Theorem 2.1. The integral

$$\int_{k/t}^\infty y^{n-1} |g^{(n)}(y)| \, dy \qquad (k > c_1 t; n = 1, 2, \cdots) \tag{7.14}$$

converges ([12], p. 251). Furthermore, the function of y defined by the integral

$$\int_{k/ty}^1 (1 - \lambda)^{-\nu} \lambda^{k+\rho-1} \, d\lambda \qquad (k/t \leqq y < \infty) \tag{7.15}$$

is monotonic increasing and bounded. Therefore the integral

$$\int_{k/t}^\infty y^{k+m} |g^{(k+m+1)}(y)| \, dy \int_{k/ty}^1 (1 - \lambda)^{-\nu} \lambda^{k+\rho-1} \, d\lambda \qquad (k > c_1 t) \tag{7.16}$$

converges ([2], p. 429). If we set $\lambda = x/y$, then the integral (7.16) assumes the form

$$\int_{k/t}^\infty |g^{(k+m+1)}(y)| \, dy \int_{k/t}^y (y - x)^{-\nu} x^{k+\rho-1} \, dx \qquad (k > c_1 t). \tag{7.17}$$

By an application of the Fubini theorem to the integral (7.17) we obtain

$$\int_{k/t}^\infty x^{k+\rho-1} \, dx \int_x^\infty (y - x)^{-\nu} |g^{(k+m+1)}(y)| \, dy \qquad (k > c_1 t). \tag{7.18}$$

Setting $x = k/u$ in (7.18), we get

$$\frac{k^\rho}{k} \int_0^t u^{-\rho} \left(\frac{k}{u}\right)^{k+1} du \int_{k/u}^\infty \left(y - \frac{k}{u}\right)^{-\nu} |g^{(k+m+1)}(y)| \, dy \qquad (k > c_1 t). \tag{7.19}$$

By Corollary 3.11, the convergence of the integral (7.13) follows from the convergence of (7.19).

We now consider the second integral on the right of (7.12). Since $\alpha(0) = \alpha(0+) = 0$, it is clear that given $\epsilon > 0$ we can determine $\eta > 0$ such that $|\alpha(y)| < \epsilon$ if $0 < y \leqq \eta$; while by Corollary 1.13, if γ is a constant greater than c_1, a constant M can be found such that $|\alpha(y)| < Me^{\gamma y}$ for $0 \leqq y < \infty$. Hence we can write

$$\begin{aligned} \left| \frac{k^{k+1}}{k!} \int_0^\infty \frac{e^{-ky/\delta}}{\delta^{k+\rho}} y^{k+\rho-1} \alpha(y) \, dy \right| &\leqq \epsilon \frac{k^{k+1}}{k!} \int_0^\eta \frac{e^{-ky/\delta}}{\delta^{k+\rho}} y^{k+\rho-1} \, dy \\ &+ M \frac{k^{k+1}}{k!} \int_\eta^\infty \frac{e^{-ky/\delta}}{\delta^{k+\rho}} e^{\gamma y} y^{k+\rho-1} \, dy \quad \left(0 < \delta < \frac{t}{2}; k > \delta\gamma\right). \end{aligned} \tag{7.20}$$

Since

$$e^{-ky/\delta} = e^{-\frac{1}{2}ky/\delta}\cdot e^{-\frac{1}{2}ky/\delta} \leqq e^{-\frac{1}{2}k\eta/\delta}e^{-\frac{1}{2}ky/\delta} \qquad (\eta \leqq y < \infty), \tag{7.21}$$

and since

$$\int_0^\eta \frac{e^{-ky/\delta}}{\delta^{k+\rho}}\, y^{k+\rho-1}\, dy < \int_0^\infty \frac{e^{-ky/\delta}}{\delta^{k+\rho}}\, y^{k+\rho-1}\, dy = \frac{\Gamma(k+\rho)}{k^{k+\rho}} \qquad (0 < \delta),$$

$$\begin{aligned} \int_\eta^\infty \frac{e^{-ky/\delta}}{\delta^{k+\rho}}\cdot e^{\gamma y} y^{k+\rho-1}\, dy &\leqq \frac{e^{-\frac{1}{2}k\eta/\delta}}{\delta^{k+\rho}} \int_\eta^\infty \exp\left[-y(\tfrac{1}{2}k/\delta - \gamma)\right] y^{k+\rho-1}\, dy \\ &< \frac{e^{-\frac{1}{2}k\eta/\delta}}{\delta^{k+\rho}} \int_0^\infty \exp\left[-y(\tfrac{1}{2}k/\delta - \gamma)\right] y^{k+\rho-1}\, dy = \frac{e^{-\frac{1}{2}k\eta/\delta}\Gamma(k+\rho)}{(\frac{1}{2}k - \gamma\delta)^{k+\rho}} \quad (k > 2\gamma\delta), \end{aligned} \tag{7.22}$$

the inequality (7.20) leads to the inequality

$$\begin{aligned} &\left| \frac{k^{k+1}}{k!} \int_0^\infty \frac{e^{-ky/\delta}}{\delta^{k+\rho}}\, y^{k+\rho-1} \alpha(y)\, dy \right| \\ &\qquad < \frac{\epsilon\Gamma(k+\rho)}{k!\,k^{\rho-1}} + \frac{e^{-\frac{1}{2}k\eta/\delta} M\Gamma(k+\rho)k^{k+1}}{(\frac{1}{2}k - \gamma\delta)^{k+\rho}\, k!} \quad (0 < \delta < \tfrac{1}{2}t;\ k > 2\gamma\delta). \end{aligned} \tag{7.23}$$

From this it follows that for any particular k we have

$$\left| \overline{\lim_{\delta\to 0+}}\ \frac{k^{k+1}}{k!} \int_0^\infty \frac{e^{-ky/\delta}}{\delta^{k+\rho}}\, y^{k+\rho-1}\alpha(y)\, dy \right| \leqq \epsilon\cdot\frac{\Gamma(k+\rho)}{k!\,k^{\rho-1}} \qquad (k > \gamma t), \tag{7.24}$$

where $\epsilon > 0$ is arbitrary. Therefore

$$\lim_{\delta\to 0+} \frac{k^{k+1}}{k!} \int_0^\infty \frac{e^{-ky/\delta}}{\delta^{k+\rho}}\, y^{k+\rho-1}\alpha(y)\, dy = 0 \qquad (k > \gamma t), \tag{7.25}$$

so that, by (7.12), we have the equation

$$\int_0^t u^{-\rho} L_{k,u}(f(x))\, du = \frac{k^{k+1}}{k!} \int_0^\infty \frac{e^{-ky/t}}{t^{k+\rho}}\, y^{k+\rho-1}\alpha(y)\, dy \qquad (k > \gamma t). \tag{7.26}$$

If we set $y = tu$, the integral on the right side of equation (7.26) becomes

$$\frac{k^{k+1}}{k!} \int_0^\infty e^{-ku} u^{k+\rho-1}\alpha(tu)\, du \qquad (k > \gamma t) \tag{7.27}$$

which can be written

$$\begin{aligned} &t^{-\rho} \frac{k^k}{(k-1)^k} \frac{(k-1)^k}{(k-1)!} \int_0^\infty e^{-(k-1)u} u^{k-1} [(tu)^\rho e^{-u}\alpha(tu)]\, du \\ &\qquad = \frac{t^{-\rho}}{(1-k^{-1})^k}\cdot L_{k-1,t}[\psi(x)] \quad (k > \gamma t), \end{aligned} \tag{7.28}$$

where

$$\psi(x) = \int_0^\infty e^{-xu}[u^\rho e^{-u/t}\alpha(u)]\, du \qquad (x > c_1), \tag{7.29}$$

t being, as it has been throughout the proof, a fixed positive number. The inversion formula for the integral (7.29) is ([11], p. 121, Theorem 3)[4]

$$\lim_{k\to\infty} L_{k-1,t}[\psi(x)] = t^{\rho} e^{-1}\alpha(t), \tag{7.30}$$

and consequently

$$\lim_{k\to\infty} \frac{t^{-\rho}}{(1-k^{-1})^k}\cdot L_{k-1,t}(\psi(x)) = \alpha(t) \tag{7.31}$$

from which, by (7.26) and (7.28), we obtain (7.1) for the particular value of t under consideration. The positive number t was, however, arbitrary, and so it is clear that (7.1) holds for $0 < t < \infty$.

8. **A second method of inversion.** Using the notation of Theorem 2.2, we have, by Corollary 3.12, the equation

$$g^{(k)}(x) = {}_{\infty}I_x^{\rho}(f^{(k)}(x)) \qquad (c < x < \infty;\ k = 1, 2, \cdots). \tag{8.1}$$

Now let the operator $L_{k,t}^{\rho}(f(x))$ be introduced by the definition

$$L_{k,t}^{\rho}(f(x)) = \frac{(-1)^k}{k!}\left(\frac{k}{t}\right)^{k+1} \{{}_{\infty}I_x^{\rho}[f^{(k)}(x)]\}_{x=k/t} \qquad (0 < t < \infty;\ k = 1, 2, \cdots). \tag{8.2}$$

By (8.1) we have

$$L_{k,t}^{\rho}(f(x)) = L_{k,t}(g(x)) \qquad (0 < t < \infty;\ k = 1, 2, \cdots), \tag{8.3}$$

and from this two theorems on inversion can be deduced immediately:

THEOREM 8.1. *If $f(x)$ is the function defined by a convergent integral of the form (2), $\alpha(t)$ being normalized, then*

$$\alpha(t) = \lim_{k\to\infty} \int_0^t L_{k,u}^{\rho}(f(x))\, du \qquad (0 < t < \infty). \tag{8.4}$$

THEOREM 8.2. *If $f(x)$ is the function defined by a convergent integral of the form (6.4), then*

$$\phi(t) = \lim_{k\to\infty} L_{k,t}^{\rho}(f(x)) \tag{8.5}$$

at all points of the Lebesgue set of $\phi(t)$.

These two theorems follow immediately from the properties ([12], p. 249; [11], p. 122) of the operator $L_{k,t}(g(x))$.

It is possible to state the theorems of Part III in terms of the operator $L_{k,t}^{\rho}(f(x))$; and in fact this method was formerly used by the author, but the methods of the present paper are simpler. Consequently, the operator $L_{k,t}^{\rho}(f(x))$ will not be used in what follows.

[4] In this case, the function $\phi(u)$ is $u^{\rho}e^{-u/t}\alpha(u)$, which clearly satisfies the conditions of the theorem.

Part III

9. **The representation theorems.** Part III will be devoted to the problem of representing functions by integrals of the form (2) and (6.4). The theorems on the possibility of such representations will be stated in terms of the operator $L_{k,t}(f(x))$. It will be found upon comparison that, if we set $\rho = 0$, the results here established reduce to those previously found ([12]; see the results tabulated on pp. 246–247) on the representation of functions by integrals of the forms (1) and (6.6). We must first, in the following section, state a theorem upon which our developments will be based.

10. **The uniqueness theorem.** We state a theorem due to D. V. Widder, who will publish a proof in a work now being prepared. For a proof of the theorem under different hypotheses on the behavior of $f(x)$ and its derivatives, see [11], p. 140. In the reference cited it will be noticed that instead of the conditions (b) and (c), the condition

$$|f^{(k)}(x)| < \frac{Mk!}{x^{k+1}} \qquad (0 < x < \infty; k = 0, 1, 2, \cdots)$$

is imposed.

THEOREM 10.1. *If the function $f(x)$ satisfies the conditions*
(a) *$f(x)$ is of class C^∞ in $(0 < x < \infty)$,*
(b) *$f^{(k)}(x) = o(x^{-k})$ $(x \to \infty; k = 0, 1, 2, \cdots)$,*
(c) *a constant $c > 0$ exists such that*

$$f^{(k)}(x) = O(e^{c/x}) \qquad (x \to 0+; k = 0, 1, 2, \cdots),$$

then $f(x)$ is given by the equation

$$f(x) = \lim_{k\to\infty} \int_0^\infty e^{-xt} L_{k,t}(f(x))\, dt \qquad (0 < x < \infty). \tag{10.1}$$

11. **Two inequalities satisfied by $L_{k,t}(f(x))$.** The inequalities established in this section will be useful in the proofs of the representation theorems.

THEOREM 11.1. *Let $f(x)$ be the function defined by a convergent integral of the form (2), and let $g(x)$ be the function defined by the corresponding integral of the form (1). Then a positive number M_0 can be found, independent of t and k, such that* (see Definition 2.1)

$$\left| \int_0^t u^{-\rho} L_{k,u}(f(x))\, du \right| \leqq M_0 \left| \int_0^\eta L_{k+m+1,w}(g(x))\, dw \right| \tag{11.1}$$

$$(0 < t < \infty; k > c_1 t),$$

where η is a positive number which does not exceed $(m + 2)t$, and

$$\int_0^t |u^{-\rho} L_{k,u}(f(x))| \, du \leqq M_0 \int_0^{(m+2)t} |L_{k+m+1,w}(g(x))| \, dw \tag{11.2}$$

$$(0 < t < \infty; k > c_1 t).$$

In the course of the proof of Theorem 7.1,[5] the existence of the integral

$$\int_0^t |u^{-\rho} L_{k,u}(f(x))| \, du \qquad (0 < t < \infty; k > c_1 t) \tag{11.3}$$

was demonstrated. Making use of (7.13)–(7.19) and Corollary 3.11, we readily obtain the equation

$$\begin{aligned} \int_0^t u^{-\rho} L_{k,u}(f(x)) \, du &= \frac{(-1)^{k+m+1}}{(k-1)! k^\rho \Gamma(1-\nu)} \\ &\quad \cdot \int_{k/t}^\infty u^{k+m} g^{(k+m+1)}(u) \, du \int_{k/(tu)}^1 (1-\lambda)^{-\nu} \lambda^{k+\rho-1} \, d\lambda \end{aligned} \tag{11.4}$$

$$(0 < t < \infty; k > c_1 t)'$$

and the inequality

$$\begin{aligned} \int_0^t |u^{-\rho} L_{k,u}(f(x))| \, du &\leqq \frac{1}{(k-1)! k^\rho \Gamma(1-\nu)} \\ &\quad \cdot \int_{k/t}^\infty u^{k+m} |g^{(k+m+1)}(u)| \, du \int_{k/(tu)}^1 (1-\lambda)^{-\nu} \lambda^{k+\rho-1} \, d\lambda \end{aligned} \tag{11.5}$$

$$(0 < t < \infty; k > c_1 t).$$

In the integrals on the right of (11.4) and (11.5), we set $u = (k + m + 1)/w$ and multiply and divide by $(k + m)!$. As a result we obtain

$$\begin{aligned} \int_0^t u^{-\rho} L_{k,u}(f(x)) \, du &= \frac{(k+m)!}{(k-1)! k^\rho \Gamma(1-\nu)} \\ &\quad \cdot \int_0^{t(k+m+1)/k} L_{k+m+1,w}(g(x)) \cdot \psi_k(\rho, w) \, dw \end{aligned} \tag{11.6}$$

$$(0 < t < \infty; k > c_1 t),$$

and

$$\begin{aligned} \int_0^t |u^{-\rho} L_{k,u}(f(x))| \, du &\leqq \frac{(k+m)!}{(k-1)! k^\rho \Gamma(1-\nu)} \\ &\quad \cdot \int_0^{t(k+m+1)/k} |L_{k+m+1,w}(g(x))| \cdot \psi_k(\rho, w) \, dw \end{aligned} \tag{11.7}$$

$$(0 < t < \infty; k > c_1 t),$$

[5] See (7.13) through (7.19). As in the proof of Theorem 7.1, c_1 is to be the greater of the two numbers c and zero.

where the function of w,

$$\psi_k(\rho, w) = \int_{kw/[t(k+m+1)]}^{1} (1 - \lambda)^{-\nu}\lambda^{k+\rho-1}\,d\lambda \quad \left(0 \leqq w \leqq \left(\frac{k+m+1}{k}\right)t\right) \tag{11.8}$$

is positive, monotonic decreasing and bounded, and such that

$$0 \leqq \psi_k(\rho, w) \leqq \psi_k(\rho, 0) = B(1 - \nu, k + \rho). \tag{11.9}$$

An application of Bonnet's theorem ([6], p. 618) to the integral on the right of equation (11.6) gives us

$$\int_0^t u^{-\rho} L_{k,u}(f(x))\,du = \frac{(k+m)!\,B(1-\nu, k+\rho)}{(k-1)!\,k^\rho\,\Gamma(1-\nu)} \int_0^\eta L_{k+m+1,w}(g(x))\,dw \qquad (0 < t < \infty;\ k > c_1 t), \tag{11.10}$$

where $0 \leqq \eta \leqq \left(\dfrac{k+m+1}{k}\right)t$. By (11.9) and the fact that $(k + m + 1)/k \leqq m + 2$, we get from (11.7) the inequality

$$\begin{aligned}&\int_0^t |u^{-\rho} L_{k,u}(f(x))|\,du \\ &\qquad \leqq \frac{(k+m)!\,B(1-\nu, k+\rho)}{(k-1)!\,k^\rho\,\Gamma(1-\nu)} \int_0^{(m+2)t} |L_{k+m+1,w}(g(x))|\,dw \qquad (0 < t < \infty;\ k > c_1 t).\end{aligned} \tag{11.11}$$

By an application of Stirling's formula we obtain

$$\frac{(k+m)!\,B(1-\nu, k+\rho)}{(k-1)!\,k^\rho\,\Gamma(1-\nu)} \leqq M_0, \tag{11.12}$$

where M_0 does not depend on k. From (11.10), (11.11) and (11.12), (11.1) and (11.2) follow directly.

12. **The representation theorems: the case where $\alpha(t)$ is bounded.** The remaining sections of Part III will be devoted to the representation theorems, that is, the theorems on the possibility of finding a representation of the form (2) for a function $f(x)$. In this section we shall consider the case where the representation is such that the function $\alpha(t)$ in the integral (2) is bounded in $(0 \leqq t < \infty)$.

CONDITION A. *A function $f(x)$ will be said to satisfy Condition* A *for the positive constant ρ if and only if*

(a) *$f(x)$ is of class C^∞ in* $(0 < x < \infty)$,
(b) $f(\infty) = 0$,
(c) *a positive constant M exists such that*

$$\left|\int_0^R t^{-\rho} L_{k,t}(f(x))\,dt\right| \leqq M \qquad (0 < R < \infty;\ k = 1, 2, \cdots),$$

(d) *a positive function $N(R)$ exists for $0 < R < \infty$ such that*

$$\int_0^R \left| t^{-\rho} L_{k,t}(f(x)) \right| dt \leqq N(R) \qquad (0 < R < \infty; k = 1, 2, \cdots).$$

The theorem that follows is stated in terms of Condition A.

THEOREM 12.1. *A necessary and sufficient condition that $f(x)$ can be represented by an integral of the form (2) in $0 < x < \infty$, where ρ is a positive constant and $\alpha(t)$ is a normalized function of bounded variation in $(0, R)$ for every positive R and is bounded in $0 \leqq t < \infty$, is that $f(x)$ should satisfy Condition A for the positive constant ρ.*

The condition will first be shown to be sufficient. In (c) set $u = k/t$, so that (c) becomes

$$\left| \int_{k/R}^{\infty} \frac{(-1)^k u^{k+\rho-1}}{k^\rho (k-1)!} f^{(k)}(u)\, du \right| \leqq M \qquad (0 < R < \infty; k = 1, 2, \cdots). \tag{12.1}$$

Making use of (12.1) and (b) of the hypothesis,[6] we can readily establish the following end-conditions:

$$\begin{aligned} &\text{(a) } f^{(k)}(x) = o(x^{-k-\rho}) \qquad (x \to \infty; k = 0, 1, 2, \cdots), \\ &\text{(b) } f^{(k)}(x) = O(x^{-k-\rho}) \qquad (x \to 0+; k = 0, 1, 2, \cdots). \end{aligned} \tag{12.2}$$

From (12.2) it is evident that $f(x)$ satisfies the hypotheses of Theorem 10.1, so that we have

$$f(x) = \lim_{k \to \infty} \int_0^\infty e^{-xt} L_{k,t}(f(x))\, dt \qquad (0 < x < \infty). \tag{12.3}$$

By (c) of the hypothesis the functions

$$\alpha_k(0) = 0, \qquad \alpha_k(t) = \int_0^t u^{-\rho} L_{k,u}(f(x))\, du \quad (0 < t < \infty; k = 1, 2, \cdots) \tag{12.4}$$

are defined for $0 \leqq t < \infty$, and

$$| \alpha_k(t) | \leqq M \qquad (0 \leqq t < \infty; k = 1, 2, \cdots), \tag{12.5}$$

while by (d) of the hypothesis, if $u_k(t)$ denotes the variation of $\alpha_k(x)$ in $(0 \leqq x \leqq t)$, we have ([6], p. 605)

$$0 \leqq u_k(t) = \int_0^t | u^{-\rho} L_{k,u}(f(x)) |\, du \leqq N(t) \quad (0 \leqq t < \infty; k = 1, 2, \cdots), \tag{12.6}$$

so that the $\alpha_k(t)$ are of uniformly bounded variation in $(0, R)$ for every positive R. It follows from (12.4) that we have

$$\int_0^\infty e^{-xt} t^\rho\, d\alpha_k(t) = \int_0^\infty e^{-xt} L_{k,t}(f(x))\, dt \qquad (0 < x < \infty; k = 1, 2, \cdots), \tag{12.7}$$

[6] The argument, which is similar to one used in [11], will not be given.

so that, by (12.3), we obtain the following equation:

$$f(x) = \lim_{k\to\infty} \int_0^\infty e^{-xt} t^\rho \, d\alpha_k(t) \qquad (0 < x < \infty). \tag{12.8}$$

Since, by (12.5) and (12.6),

$$\begin{aligned} |\alpha_k(t)| &\leqq M \qquad (0 \leqq t \leqq 1; k = 1, 2, \cdots), \\ u_k(1) &\leqq N(1) \qquad (k = 1, 2, \cdots), \end{aligned} \tag{12.9}$$

it is possible ([5], p. 265) to find a subsequence $\{\alpha_{1,k}(t)\}$ of the sequence $\{\alpha_k(t)\}$ such that $\lim_{k\to\infty} \alpha_{1,k}(t)$ exists for $0 \leqq t \leqq 1$; the limit-function does not exceed M in absolute value, and is of total variation not exceeding $N(1)$ in $0 \leqq t \leqq 1$. Now consider the interval $(0, 2)$. By the previous argument, we can find a subsequence $\{\alpha_{2,k}(t)\}$ of the sequence $\{\alpha_{1,k}(t)\}$ such that $\lim_{k\to\infty} \alpha_{2,k}(t)$ exists for $0 \leqq t \leqq 2$; the limit-function does not exceed M in absolute value, and is of total variation not exceeding $N(2)$ in $0 \leqq t \leqq 2$. Proceeding in this way, for every positive integer n we can find a sequence $\{\alpha_{n,k}(t)\}$ which is a subsequence of $\{\alpha_{n-1,k}(t)\}$ and such that $\lim_{k\to\infty} \alpha_{n,k}(t)$ exists for $0 \leqq t \leqq n$; the limit-function does not exceed M in absolute value, and is of total variation not exceeding $N(n)$ in $0 \leqq t \leqq n$. The sequence $\{\alpha_{k,k}(t)\}$ therefore approaches a limit $\alpha(t)$ for all non-negative values of t, and $\alpha(t)$ is of bounded variation in $(0, R)$ for every positive R and is such that

$$|\alpha(t)| \leqq M \qquad (0 \leqq t < \infty). \tag{12.10}$$

It is furthermore clear that $\alpha(0) = 0$.

Now let x be a fixed positive number, and let $\epsilon > 0$ be given. For $R > \rho/x$ the function $e^{-xR}R^\rho$ is monotonic decreasing, and so, by the second law of the mean for Stieltjes integrals (cf. [1]) and (12.5), we obtain

$$\begin{gathered} \left| \int_R^\infty e^{-xt} t^\rho \, d\alpha_{k,k}(t) \right| \leqq e^{-xR} R^\rho \sup_{R \leqq S < \infty} \left| \int_R^S d\alpha_{k,k}(t) \right| \leqq 2M e^{-xR} R^\rho \\ \left(R > \frac{\rho}{x}; k = 1, 2, \cdots \right), \end{gathered} \tag{12.11}$$

so that R_0 can be found such that

$$\left| \int_R^\infty e^{-xt} t^\rho \, d\alpha_{k,k}(t) \right| < \epsilon \qquad (R > R_0; k = 1, 2, \cdots). \tag{12.12}$$

Now let R be a number greater than R_0. Since the functions $\alpha_{k,k}(t)$ are uniformly bounded and uniformly of bounded variation in $(0, R)$, we obtain, by an application of the Helly-Bray theorem ([4], p. 15), the equation

$$\lim_{k\to\infty} \int_0^R e^{-xt} t^\rho \, d\alpha_{k,k}(t) = \int_0^R e^{-xt} t^\rho \, d\alpha(t). \tag{12.13}$$

We now write

$$\int_0^\infty e^{-xt}t^\rho\,d\alpha_{k,k}(t) - \int_0^R e^{-xt}t^\rho\,d\alpha_{k,k}(t) = \int_R^\infty e^{-xt}t^\rho\,d\alpha_{k,k}(t) \tag{12.14}$$

$$(0 < R < \infty;\, k = 1, 2, \cdots),$$

so that, for $R > R_0$, we have

$$\left| \int_0^\infty e^{-xt}t^\rho\,d\alpha_{k,k}(t) - \int_0^R e^{-xt}t^\rho\,d\alpha_{k,k}(t) \right| < \epsilon \tag{12.15}$$

$$(R_0 < R < \infty;\, k = 1, 2, \cdots).$$

Now holding R fast, we let k become infinite in (12.15). By (12.8) and (12.13) we obtain

$$\left| f(x) - \int_0^R e^{-xt}t^\rho\,d\alpha(t) \right| \leqq \epsilon \tag{12.16}$$

for every $R > R_0$. It follows that

$$f(x) = \lim_{R\to\infty} \int_0^R e^{-xt}t^\rho\,d\alpha(t) = \int_0^\infty e^{-xt}t^\rho\,d\alpha(t). \tag{12.17}$$

Since x is an arbitrary positive number, it follows that (12.17) holds for all positive x; we have $\alpha(0) = 0$, and we set $\alpha(0+) = 0$ and $\alpha(t) = \frac{1}{2}[\alpha(t+) + \alpha(t-)]$, as we may without changing the value of $f(x)$, so that $\alpha(t)$ is normalized.

Now suppose that $f(x)$ has the representation, and define $g(x)$ as in Theorem 2.1. Obviously $f(x)$ satisfies (a) of Condition A, while by Theorem 4.1 it is clear that $f(x)$ satisfies (b) also. Furthermore, it is known that a constant M_1 and a positive function $N_1(R)$ exist such that ([12], p. 272)

$$\begin{aligned} &\left| \int_0^R L_{k,t}(g(x))\,dt \right| \leqq M_1 && (0 < R < \infty;\, k = 1, 2, \cdots), \\ &\int_0^R |\, L_{k,t}(g(x))\,|\,dt \leqq N_1(R) && (0 < R < \infty;\, k = 1, 2, \cdots). \end{aligned} \tag{12.18}$$

Hence, by Theorem 11.1, we obtain the inequalities

$$\left| \int_0^R t^{-\rho} L_{k,t}(f(x))\,dt \right| \leqq M_0 \left| \int_0^\eta L_{k+m+1,t}(g(x))\,dt \right| \leqq M_0 \cdot M_1 = M$$

$$(0 < R < \infty;\, k = 1, 2, \cdots),$$

$$\begin{aligned} \int_0^R |\, t^{-\rho} L_{k,t}(f(x))\,|\,dt &\leqq M_0 \int_0^{(m+2)R} |\, L_{k+m+1,t}(g(x))\,|\,dt \\ &\leqq M_0 \cdot N_1((m+2)R) = N(R) \end{aligned} \tag{12.19}$$

$$(0 < R < \infty;\, k = 1, 2, \cdots).$$

Condition A has therefore been shown to be necessary.

13. $\alpha(t)$ **of bounded variation in** $(0 \leqq t < \infty)$.

CONDITION B. *A function $f(x)$ will be said to satisfy Condition* B *for the positive constant ρ if and only if $f(x)$ satisfies Condition* A *for the positive constant ρ with the function $N(R)$ bounded.*

THEOREM 13.1. *A necessary and sufficient condition that $f(x)$ can be represented in $(0 < x < \infty)$ by an integral of the form* (2), *where $\alpha(t)$ is a normalized function of bounded variation in $(0 \leqq t < \infty)$ and ρ is a positive constant, is that $f(x)$ should satisfy Condition* B *for the positive constant ρ.*

Theorem 13.1 follows readily from Theorem 12.1. The proof will therefore be omitted.

14. **The general case.** In the present section we deal with the case where $\alpha(t)$ is merely of bounded variation in $(0, R)$ for every positive R.

THEOREM 14.1. *A necessary and sufficient condition that $f(x)$ can be represented in $(0 < x < \infty)$ by an integral of the form* (2), *where $\alpha(t)$ is of bounded variation in $(0, R)$ for every positive R and ρ is a positive constant, is that, for every $\epsilon > 0$, $f(x + \epsilon)$ should satisfy Condition* A *for the positive constant ρ.*

Suppose $f(x)$ satisfies the condition. Then for $\epsilon > 0$ we have

$$f(x + \epsilon) = \int_0^\infty e^{-xt} t^\rho \, d\alpha_\epsilon(t) \qquad (0 < x < \infty) \tag{14.1}$$

by Theorem 12.1, where $\alpha_\epsilon(t)$ is a normalized function of bounded variation in $(0, R)$ for every positive R and is bounded in $(0 \leqq t < \infty)$. Then by replacing $x + \epsilon$ by x we obtain the representation

$$f(x) = \int_0^\infty e^{-xt} t^\rho \, d\alpha(t) \qquad (\epsilon < x < \infty), \tag{14.2}$$

where we have set

$$\alpha(0) = 0, \qquad \alpha(t) = \int_0^t e^{\epsilon u} \, d\alpha_\epsilon(u) \qquad (0 < t < \infty), \tag{14.3}$$

the function $\alpha(t)$ clearly being normalized. By Theorem 5.1 we see that $\alpha(t)$ is unique, and hence does not depend on ϵ; and since $\epsilon > 0$ is arbitrary, it is clear that (14.2) must hold for $(0 < x < \infty)$.

Now suppose that $f(x)$ has the representation. For each $\epsilon > 0$ we have

$$f(x + \epsilon) = \int_0^\infty e^{-xt} t^\rho \, d\alpha_\epsilon(t) \qquad (-\epsilon < x < \infty), \tag{14.4}$$

where

$$\alpha_\epsilon(0) = 0, \qquad \alpha_\epsilon(t) = \int_0^t e^{-\epsilon u} \, d\alpha(u) \qquad (0 < t < \infty). \tag{14.5}$$

By Corollary 1.13 we can find a constant K_ϵ such that $|\alpha(t)| < K_\epsilon e^{\frac{1}{2}\epsilon t}$ for $(0 \leqq t < \infty)$. By Stieltjes integration by parts we get

$$\alpha_\epsilon(t) = e^{-\epsilon t}\alpha(t) + \epsilon \int_0^t e^{-\epsilon t}\alpha(t)\,dt \qquad (0 < t < \infty). \tag{14.6}$$

From this we readily obtain the inequality

$$|\alpha_\epsilon(t)| < K_\epsilon e^{-\frac{1}{2}\epsilon t} + K_\epsilon \cdot \epsilon \int_0^t e^{-\frac{1}{2}\epsilon t}\,dt \qquad (0 < t < \infty) \tag{14.7}$$

which becomes

$$|\alpha_\epsilon(t)| < K_\epsilon e^{-\frac{1}{2}\epsilon t} + 2K_\epsilon(1 - e^{-\frac{1}{2}\epsilon t}) \qquad (0 < t < \infty), \tag{14.8}$$

so that $\alpha_\epsilon(t)$ is bounded in $(0 \leqq t < \infty)$. Hence, by Theorem 13.1, it follows that for each $\epsilon > 0$ the inequalities (c) and (d) of Condition A are satisfied by the function $f(x + \epsilon)$. Since (a) and (b) of Condition A are obviously satisfied by $f(x + \epsilon)$, it has been shown that the condition is necessary.

15. **Completely monotonic functions.** A function $f(x)$ is said to be completely monotonic in $(0 < x < \infty)$ if

$$(-1)^k f^{(k)}(x) \geqq 0 \qquad (k = 0, 1, 2, \cdots; 0 < x < \infty). \tag{15.1}$$

Our next two theorems will deal with the representation of completely monotonic functions by integrals of the form (2). The theorems will be stated in terms of the following conditions:

CONDITION C (C′). *A function $f(x)$ will be said to satisfy Condition* C (C′) *for the positive constant ρ if and only if*

(a), (a′) *$f(x)$ is completely monotonic in* $(0 < x < \infty)$,

(b), (b′) $f(\infty) = 0$,

(c) *the integral*

$$\int_0^\infty x^{\rho-1} f(x)\,dx$$

is convergent,

(c′) *the integral*

$$\int_1^\infty x^{\rho-1} f(x)\,dx$$

is convergent.

LEMMA 15.1. *Let $f(x)$ be of class C^1 in $(0 < x < \infty)$; suppose that $f(x) \geqq 0$ and $f'(x) \leqq 0$, and that the integral*

$$\int_0^\infty x^{a-1} f(x)\,dx \tag{15.2}$$

converges, a being a positive constant. Then the following equation holds:

$$\int_0^\infty x^{a-1}f(x)\,dx = -\frac{1}{a}\int_0^\infty x^a f'(x)\,dx. \tag{15.3}$$

We have, if $A > 0$, the inequality

$$\int_{\frac{1}{2}A}^{A} x^{a-1}f(x)\,dx \geqq f(A)\int_{\frac{1}{2}A}^{A} x^{a-1}\,dx \geqq f(A)A^a\left[\frac{2^a-1}{a\cdot 2^a}\right] \geqq 0. \tag{15.4}$$

If, in (15.4), we allow A to become infinite, or to approach zero, the integral on the left will approach zero, since the integral (15.2) converges. Hence we have

$$\lim_{x\to 0+} x^a f(x) = \lim_{x\to\infty} x^a f(x) = 0. \tag{15.5}$$

Now suppose that $0 < \alpha < \beta < \infty$; by integration by parts we obtain

$$\int_\alpha^\beta x^{a-1}f(x)\,dx = \frac{\beta^a f(\beta)}{a} - \frac{\alpha^a f(\alpha)}{a} - \frac{1}{a}\int_\alpha^\beta x^a f'(x)\,dx. \tag{15.6}$$

In (15.6) we allow α to approach zero, and β to become infinite. By (15.5) and the convergence of the integral (15.2), we see that the integral in the right side of (15.6) approaches a limit, and that (15.3) holds.

LEMMA 15.2. *Let $f(x)$ be of class C^1 in $(0 < x < \infty)$. Suppose that $f(x) \geqq 0$ and $f'(x) \leqq 0$, that $\lim_{x\to\infty} x^a f(x)$ exists, and that the integral*

$$\int_0^\infty x^a f'(x)\,dx \tag{15.7}$$

converges, a being a positive constant. Then equation (15.3) *holds.*

In this case also, we can write equation (15.6), and from it obtain

$$-\frac{1}{a}\int_\alpha^\beta x^a f'(x)\,dx = \frac{\alpha^a f(\alpha)}{a} - \frac{\beta^a f(\beta)}{a} + \int_\alpha^\beta x^{a-1}f(x)\,dx. \tag{15.8}$$

By the convergence of the integral (15.7) the right side of (15.8) approaches a limit as α approaches zero. The integral in the right side of (15.8) is clearly non-negative and non-decreasing as α decreases, and so must either become infinite or approach a non-negative limit as α approaches zero, while by hypothesis $\alpha^a f(\alpha) \geqq 0$. Hence it is clear that

$$\lim_{\alpha\to 0+}\int_\alpha^\beta x^{a-1}f(x)\,dx \tag{15.9}$$

exists. It then follows that $\lim_{\alpha\to 0+} \alpha^a f(\alpha)$ exists also, and we have

$$-\frac{1}{a}\int_0^\beta x^a f'(x)\,dx = \frac{1}{a}\lim_{\alpha\to 0+}\alpha^a f(\alpha) - \frac{\beta^a f(\beta)}{a} + \int_0^\beta x^{a-1}f(x)\,dx. \tag{15.10}$$

In (15.10) we allow β to become infinite. From the convergence of the integral (15.7) and the existence of $\lim_{\beta\to\infty} \beta^a f(\beta)$ it is clear that the integral in the right side of (15.10) must approach a limit, and so we have

$$-\frac{1}{a}\int_0^\infty x^a f'(x)\,dx = \frac{1}{a}\,[\lim_{\alpha\to0+} \alpha^a f(\alpha) - \lim_{\beta\to\infty}\beta^a f(\beta)] + \int_0^\infty x^{a-1}f(x)\,dx. \tag{15.11}$$

We must show that both $\lim_{x\to0+} x^a f(x)$ and $\lim_{x\to\infty} x^a f(x)$ are zero. Suppose that $\lim_{x\to\infty} x^a f(x) = A > 0$. Then a positive number x_0 exists such that $x^{a-1}f(x) \geqq \frac{1}{2}A/x$ for $x_0 \leqq x$, and this implies the divergence of the integral in the right side of (15.11). Hence $\lim_{x\to\infty} x^a f(x)$, which is evidently non-negative, must be zero. That $\lim_{x\to0+} x^a f(x) = 0$ is shown similarly, so that (15.3) follows from (15.11).

THEOREM 15.3. *A necessary and sufficient condition that $f(x)$ can be represented in $(0 < x < \infty)$ by an integral of the form (2) where ρ is a positive constant and $\alpha(t)$ is non-decreasing and bounded in $(0 \leqq t < \infty)$ is that $f(x)$ should satisfy Condition C for the positive constant ρ.*

Starting from (c) of the hypothesis, we obtain, by making k successive applications of Lemma 15.1 with $a = \rho$, the equation

$$\int_0^\infty x^{\rho-1}f(x)\,dx = \frac{1}{\rho(\rho+1)\cdots(\rho+k-1)}\int_0^\infty x^{k+\rho-1}(-1)^k f^{(k)}(x)\,dx \tag{15.12}$$

$$(k = 1, 2, \cdots).$$

In the integral on the right of equation (15.12) we set $x = k/t$, then we divide both sides of the equation by $k^\rho(k-1)!$ and multiply both sides by $\Gamma(k+\rho)/\Gamma(\rho)$; since

$$\frac{\Gamma(k+\rho)}{\Gamma(\rho)} = \rho(\rho+1)\cdots(\rho+k-1) \qquad (k = 1, 2, \cdots), \tag{15.13}$$

the result of these operations is the equation

$$\frac{\Gamma(k+\rho)}{(k-1)!\,k^\rho\,\Gamma(\rho)}\int_0^\infty x^{\rho-1}f(x)\,dx = \int_0^\infty t^{-\rho}L_{k,t}(f(x))\,dt \qquad (k = 1, 2, \cdots). \tag{15.14}$$

Since for a completely monotonic function we must have

$$L_{k,t}(f(x)) = |\,L_{k,t}(f(x))\,| \qquad (0 < t < \infty;\ k = 1, 2, \cdots), \tag{15.15}$$

and since by (11.12) we can find a constant $M_0 > 0$ such that

$$\begin{aligned}\frac{\Gamma(k+\rho)}{k^\rho(k-1)!} &= \frac{(k+m)!\,\Gamma(1-\nu)\Gamma(k+\rho)}{(k-1)!\,k^\rho\,\Gamma(1-\nu)\Gamma(k+m+1)}\\ &= \frac{(k+m)!\,B(1-\nu, k+\rho)}{(k-1)!\,k^\rho\,\Gamma(1-\nu)} \leqq M_0 \quad (k = 1, 2, \cdots),\end{aligned} \tag{15.16}$$

we derive from (15.14) the inequality

$$\int_0^\infty |t^{-\rho} L_{k,t}(f(x))|\, dt \leqq \frac{M_0}{\Gamma(\rho)} \int_0^\infty x^{\rho-1} f(x)\, dx = M_1 \qquad (k = 1, 2, \cdots). \tag{15.17}$$

It follows from Theorem 13.1 that $f(x)$ has the representation

$$f(x) = \int_0^\infty e^{-xt} t^\rho \, d\alpha(t) \qquad (0 < x < \infty), \tag{15.18}$$

where $\alpha(t)$ is a normalized function of bounded variation in $(0 \leqq t < \infty)$. Since by Theorem 7.1 we may write

$$\alpha(t_2) - \alpha(t_1) = \lim_{k\to\infty} \int_{t_1}^{t_2} t^{-\rho} L_{k,t}(f(x))\, dt \quad (0 < t_1 < t_2 < \infty), \tag{15.19}$$

it is clear, by (15.15), that

$$\alpha(t_2) - \alpha(t_1) \geqq 0 \qquad (0 < t_1 < t_2 < \infty). \tag{15.20}$$

From Theorem 7.1 and (15.15) it follows that $\alpha(t) \geqq 0$, so that, since $\alpha(0) = 0$, (15.20) holds for $(0 \leqq t_1 < t_2 < \infty)$. Hence $\alpha(t)$ is non-decreasing and bounded in $(0 \leqq t < \infty)$ as the theorem asserts.

Now suppose that $f(x)$ has the representation. By Theorem 3.1 we have

$$(-1)^k f^{(k)}(x) = \int_0^\infty e^{-xt} t^{k+\rho}\, d\alpha(t) \geqq 0 \qquad (0 < x < \infty;\, k = 1, 2, \cdots), \tag{15.21}$$

since $\alpha(t)$ is non-decreasing, so that (a) of Condition C is satisfied. By Theorem 4.1 it is clear that (b) of Condition C is also satisfied.

Since $\alpha(t)$ is bounded and non-decreasing in $(0 \leqq t < \infty)$, it is of bounded variation there, and so, by Theorem 13.1, a constant K_0 exists such that

$$\int_0^R |t^{-\rho} L_{k,t}(f(x))|\, dt \leqq K_0 \qquad (0 < R < \infty;\, k = 1, 2, \cdots). \tag{15.22}$$

Since the integral on the left of (15.22) is a bounded, non-decreasing function of R, we infer the convergence of the integral

$$\int_0^\infty |t^{-\rho} L_{k,t}(f(x))|\, dt \qquad (k = 1, 2, \cdots). \tag{15.23}$$

We now consider the integral (15.23) written for $k = 1$. Remembering that $-f'(x) \geqq 0$, and setting $t = x^{-1}$, we obtain the convergent integral

$$\int_0^\infty x^\rho [-f'(x)]\, dx. \tag{15.24}$$

By Theorem 4.1, we see that $\lim_{x\to\infty} x^\rho f(x) = 0$. Hence, by Lemma 15.2, we conclude that (c) of Condition C is satisfied. This completes the proof of the necessity of the condition.

THEOREM 15.4. *A necessary and sufficient condition that $f(x)$ can be represented in $(0 < x < \infty)$ by an integral of the form (2) where ρ is a positive constant and $\alpha(t)$ is non-decreasing in $(0 \leqq t < \infty)$ is that $f(x)$ should satisfy Condition C′ for the positive constant ρ.*

From (c′) of the hypothesis, it is evident that the integral

$$\int_{\eta}^{\infty} x^{\rho-1} f(x)\, dx \tag{15.25}$$

converges, where η is an arbitrary positive number. Suppose $h > 0$; then the function $[(x - \eta)/x]^{\rho-1}$ is bounded and monotonic in $(\eta + h \leqq x < \infty)$, being increasing if $1 < \rho$ and decreasing if $0 < \rho < 1$. If the integrand of the convergent integral

$$\int_{\eta+h}^{\infty} x^{\rho-1} f(x)\, dx \tag{15.26}$$

is multiplied by $[(x - \eta)/x]^{\rho-1}$, the resulting integral will converge ([2], p. 429). Furthermore, the integral

$$\int_{\eta}^{\eta+h} (x - \eta)^{\rho-1} f(x)\, dx \tag{15.27}$$

obviously converges since $0 < \rho$ and $f(x)$ is continuous in $(0 < x < \infty)$. Hence the integral

$$\int_{\eta}^{\infty} (x - \eta)^{\rho-1} f(x)\, dx \tag{15.28}$$

converges.[7] In (15.28) we set $y = x - \eta$ and $\phi_\eta(y) = f(y + \eta)$, and so obtain the convergent integral

$$\int_0^{\infty} y^{\rho-1} \phi_\eta(y)\, dy. \tag{15.29}$$

From the definition of $\phi_\eta(y)$ we get

$$(-1)^k \phi_\eta^{(k)}(y) = (-1)^k f^{(k)}(y + \eta) \geqq 0 \quad (-\eta < y < \infty; k = 0, 1, \cdots), \tag{15.30}$$

so that $\phi_\eta(y)$ is completely monotonic in $(-\eta < y < \infty)$, while obviously $\phi_\eta(\infty) = f(\infty) = 0$. Hence $\phi_\eta(y)$ satisfies Condition C, and by Theorem 15.3 we have

$$\phi_\eta(y) = \int_0^{\infty} e^{-yt} t^{\rho}\, d\beta_\eta(t) \qquad (0 < y < \infty), \tag{15.31}$$

where $\beta_\eta(t)$ is bounded and non-decreasing in $(0 \leqq t < \infty)$. In (15.31) we set $y = x - \eta$ and

$$\alpha(0) = 0, \qquad \alpha(t) = \int_0^{t} e^{\eta t}\, d\beta_\eta(t) \qquad (0 < t < \infty), \tag{15.32}$$

[7] Note that we have actually proved this only if $\rho \neq 1$; but for $\rho = 1$ the integrals (15.28) and (15.25) are identical.

and obtain

$$f(x) = \int_0^\infty e^{-xt} t^\rho \, d\alpha(t) \qquad (\eta < x < \infty). \tag{15.33}$$

Since $\beta_\eta(t)$ is normalized, it is clear that $\alpha(t)$ is normalized, and $\alpha(t)$ is also obviously non-decreasing. By Theorem 5.1 the function $\alpha(t)$ is independent of η, and so, since η is arbitrary, the representation (15.33) must be valid for $(0 < x < \infty)$.

Now suppose that $f(x)$ has the representation. Since $\alpha(t)$ is non-decreasing in $(0 \leqq t < \infty)$, it is clear that $f(x)$ is completely monotonic in $(0 < x < \infty)$, so that (a′) of Condition C′ is satisfied, while obviously (b′) is satisfied also.

If we set $\phi(x) = f(x + 1)$, we get

$$\phi(x) = \int_0^\infty e^{-xt} t^\rho \, d\alpha_1(t) \qquad (-1 < x < \infty), \tag{15.34}$$

where the function $\alpha_1(t)$ is defined as follows:

$$\alpha_1(0) = 0, \qquad \alpha_1(t) = \int_0^t e^{-u} \, d\alpha(u) \qquad (0 < t < \infty). \tag{15.35}$$

Since by hypothesis $f(1)$ exists, it follows from Theorem 1.1 that

$$\int_0^\infty e^{-u} \, d\alpha(u) \tag{15.36}$$

converges, and hence that $\lim_{t\to\infty} \alpha_1(t)$ exists, so that $\alpha_1(t)$, which by its definition is non-decreasing, must be bounded in $(0 \leqq t < \infty)$. The function $\phi(x)$ therefore satisfies Condition C, so that

$$\int_0^\infty x^{\rho-1} \phi(x) \, dx \tag{15.37}$$

must converge. If in (15.37) we replace $\phi(x)$ by $f(x + 1)$ and set $y = x + 1$, we obtain the convergent integral

$$\int_1^\infty (y - 1)^{\rho-1} f(y) \, dy. \tag{15.38}$$

We break (15.38) up into the sum of two integrals taken over the intervals $(1, 1 + h)$ and $(1 + h, \infty)$ respectively, h being positive. Multiplication of the integrand of the second of these by the monotonic bounded function $[y/(y - 1)]^{\rho-1}$ gives the convergent integral

$$\int_{1+h}^\infty y^{\rho-1} f(y) \, dy, \tag{15.39}$$

while the existence of

$$\int_1^{1+h} y^{\rho-1} f(y) \, dy \tag{15.40}$$

is assured by the continuity of $f(y)$ in $(1 \leqq y \leqq 1 + h)$. Therefore the integral (c′) of Condition C′ has been proved convergent, and the necessity of the condition has been established.

16. $\alpha(t)$ **absolutely continuous.** Our remaining results concern the possibility of representing functions by integrals of the form (6.4). The proofs will not be given, as they offer no difficulty and can readily be supplied by the reader. For the sake of brevity, we shall summarize the theorems in a table. In each case it must be remembered that $f(x)$ is of class C^∞ in $(0 < x < \infty)$ and that $f(\infty) = 0$, that ρ is a positive constant, that $\phi(t)$ is of class $L(0, R)$ for every positive R, and that the given condition is necessary and sufficient for the representation of $f(x)$ by an integral of the form (6.4) with $\phi(t)$ as described.

THEOREM 16.1. *$|\phi(t)| \leqq M$ in $(0 \leqq t < \infty)$, M being a constant;*

$$| t^{-\rho} L_{k,t}(f(x)) | \leqq M \qquad (0 < t < \infty;\ k = 1, 2, \cdots).$$

THEOREM 16.2. *$\phi(t)$ is of class $L(0, \infty)$;*

$$\lim_{\substack{m\to\infty \\ n\to\infty}} \int_0^\infty t^{-\rho} | L_{m,t}(f(x)) - L_{n,t}(f(x)) | \, dt = 0.$$

THEOREM 16.3. *$\phi(t)$ is of class $L^p(0, \infty)$, where p is a constant and $p > 1$;*

$$\int_0^\infty | t^{-\rho} L_{k,t}(f(x)) |^p \, dt \leqq M \qquad (M \text{ a constant};\ k = 1, 2, \cdots).$$

THEOREM 16.4. *$\phi(t)$ is of class $L(0, R)$ for every positive R; for every positive ϵ, $f(x + \epsilon)$ satisfies Condition* A, *and the function defined by*

$$\lim_{k\to\infty} \int_0^t u^{-\rho} L_{k,u}(f(x)) \, du,$$

which exists for every positive t, is absolutely continuous.

BIBLIOGRAPHY

1. R. P. BOAS AND D. V. WIDDER, *The iterated Stieltjes transform*, Transactions of the American Mathematical Society, vol. 45(1939), pp. 1–72.
2. T. J. I'A. BROMWICH, *An Introduction to the Theory of Infinite Series*, 1908.
3. T. CARLEMAN, *Sur les Equations Intégrales Singulières à Noyau Rèel et Symétrique*, 1912.
4. G. C. EVANS, *The Logarithmic Potential, Discontinuous Dirichlet and Neumann Problems*, American Mathematical Society Colloquium Publications, vol. 6, 1927.
5. E. HELLY, *Über lineare Funktionaloperationen*, Sitzungsberichte der mathematisch-naturwissenschaftlichen Klasse der kaiserlichen Akademie, Vienna, vol. 121(1912).
6. E. W. HOBSON, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, vol. 1, 1927.
7. E. W. HOBSON, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, vol. 2, 1926.
8. E. C. TITCHMARSH, *The Theory of Functions*, 1932.

9. D. V. WIDDER, *A generalization of Dirichlet's series and of Laplace's integrals by means of a Stieltjes integral*, Transactions of the American Mathematical Society, vol. 31(1929), pp. 694–743.
10. D. V. WIDDER, *Necessary and sufficient conditions for the representation of a function as a Laplace integral*, Transactions of the American Mathematical Society, vol. 33(1931), pp. 851–892.
11. D. V. WIDDER, *The inversion of the Laplace integral and the related moment problem*, Transactions of the American Mathematical Society, vol. 36(1934), pp. 107–200.
12. D. V. WIDDER, *A classification of generating functions*, Transactions of the American Mathematical Society, vol. 39(1936), pp. 244–298.
13. A. ZYGMUND, *Trigonometrical Series*, 1935.

LAFAYETTE COLLEGE.

PROJECTIONS IN MINKOWSKI AND BANACH SPACES

By Andrew Sobczyk

Introduction. In his now classic *Théorie des Opérations Linéaires* Banach proposed the following problem: Given any closed linear subspace of a Lebesgue function space L_p (or of a sequence space l_p), $1 < p \neq 2$, does there always exist a complementary closed linear subspace? Or, equivalently, does there always exist a *projection* on any closed linear subspace of L_p or of l_p? The question has recently been answered, in the negative, by F. J. Murray [6].[1]

Bohnenblust [3] investigated projections of n-dimensional Minkowski spaces on $(n-1)$-dimensional subspaces, with a view toward illuminating the question of the existence of projections in general Banach spaces. In this paper we take further steps in this direction.[2]

We first obtain, after necessary preliminaries to later general considerations (§1), in §2 the results of Murray by a briefer method, and in addition quantitative information which Murray did not obtain. In §3 we discuss orthogonal projections, and apply the results to obtain further quantitative information. Various generalizations of l_p-spaces are then introduced. In §4, we study a class of Banach spaces S, of which the elements are infinite sequences $x = \{x_i\}$, and which have the following symmetry property: If $x = \{x_i\}$ is any element of S, then $\{|x_i|\}$ is also an element of S, and $\|\{x_i\}\| = \|\{|x_i|\}\|$. These spaces include Banach spaces with a base $\{X_i\}$ having the corresponding symmetry property: if $x = \sum_{i=1}^{\infty} x_i X_i$ is the expansion of an element, then $\sum_{i=1}^{\infty} |x_i| \cdot X_i$ is an element, and $\|x\| = \|\sum_{i=1}^{\infty} |x_i| \cdot X_i\|$. In any space S, a Euclidean norm $\|x\|_2$ is introduced on a certain dense linear subset, and it is shown that if a projection exists for every closed linear subspace, then the Euclidean radii of the unit sphere of S in certain directions must be bounded both from 0 and from ∞. In particular, if for a space S these directions are "minimal" or "maximal", this is sufficient to require the space to be isomorphic to Hilbert space.

In §5, we study a type of spaces S which are generated by two-dimensional norms, in particular, spaces defined by a sequence $p_2, p_3, p_4, \cdots$ of exponents. These spaces specialize to l_p-spaces in case $p = p_2 = p_3 = \cdots$. Finally, in §6,

Received May 23, 1940; presented to the American Mathematical Society, September 5, 1939.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

[2] The writer wishes to acknowledge his indebtedness to Professor Bohnenblust, both for suggesting the original problem of this investigation, and for stimulation and help received in our many discussions.

we study Orlicz spaces, or spaces of sequences in which the norm is defined by a monotone function $M(t)$. An Orlicz space specializes to an l_p-space in case $M(t) = | t |^p$. For general $p_2p_3p_4 \cdots$ and Orlicz spaces, it is shown likewise that if a projection exists for every closed linear subspace, the space must be isomorphic to Hilbert space.

Our results strongly suggest the following general conjecture: A Banach space, such that there exists a projection on every closed linear subspace, the norms of the projections being uniformly bounded, is isomorphic to a Hilbert space.

Throughout we have restricted our attention to the case of real linear spaces; no essential changes are necessary to extend all of the results to the case of complex linear subspaces of complex linear spaces.

We now recall to the reader certain definitions and a theorem found in Banach's book [2], and explain some conventions of our terminology and notation.

A transformation or operation T, defined for every element of a normed linear space, is called *linear* if and only if it is additive, homogeneous, and continuous (= bounded). Two normed linear spaces are *isomorphic* ([2], p. 180) if there exists a one-to-one transformation T between the spaces which is linear in both directions. The spaces are *equivalent* if in addition $| T | = | T^{-1} | = 1$.

A Banach space is any complete normed linear space. It follows from Theorem 5, p. 41, of Banach [2] that for two Banach spaces to be isomorphic, it is sufficient that the transformation T be one-to-one and linear in one direction.

Two norms $\| x \|$ and $\| x \|_1$ on the same linear space are isomorphic if for all x

$$C_1 \cdot \| x \| \leqq \| x \|_1 \leqq C_2 \cdot \| x \|,$$

where $C_1 > 0$, $C_2 > 0$. If any two norms on the same linear space are isomorphic, the topologies defined by the two norms are equivalent. By a *Banach norm* on a linear space we mean of course a norm which makes the space a Banach space. By the theorem of Banach already mentioned, it follows that in order that two Banach norms $\| x \|$ and $\| x \|_1$ on the same linear space be isomorphic, it is sufficient that there exist $C > 0$ such that the inequality

$$C \cdot \| x \| \leqq \| x \|_1$$

is satisfied for all x. We shall have occasion to use this fact in §4.

A *Minkowski space* is any finite-dimensional normed linear space.[3] Corresponding to any choice of linearly independent elements of the space, a system of coördinates may be introduced. The norm in any Minkowski space is isomorphic with the Euclidean norm of any such coördinate system. By the Bolzano-Weierstrass theorem, a Minkowski space is of course complete and so a Banach space.

General Banach spaces will usually be denoted by B or L, and closed linear

[3] Historically, Minkowski was led to introduce these spaces by his researches on quadratic forms and the theory of numbers. Banach's generalization to infinite dimensionality came some years later.

subspaces by B_1 or l. A general n-dimensional Minkowski space will be denoted by l_n. The following notations are the same as those used by Murray in [6], with extensions to include additional cases which we wish to consider. The Minkowski spaces of sequences $x = (x_1, \cdots, x_n)$, with the norm defined by

$$\| x \| = \left(\sum_1^n | x_i |^p\right)^{1/p}, \qquad 1 \leqq p < \infty,$$

$$\| x \| = \max_i | x_i |, \qquad p = \infty,$$

will be denoted by $l_{p,n}$ and $l_{\infty,n}$ respectively. The Banach spaces of infinite sequences $x = \{x_i\}$ such that $\sum_1^\infty | x_i |^p$ is convergent, with norm

$$\| x \| = \left(\sum_1^\infty | x_i |^p\right)^{1/p}, \qquad 1 \leqq p < \infty,$$

will be designated as $l_p = l_{p,\infty}$; and the Banach space (m) of all bounded sequences $x = \{x_i\}$, with norm

$$\| x \| = \text{l.u.b.}_i | x_i |, \qquad p = \infty,$$

as $l_\infty = l_{\infty,\infty}$. By L_p and L_∞ we shall mean respectively the Banach space of measurable functions[4] $x(t)$ on a finite or infinite interval such that $\int | x(t) |^p \, dt$ is finite, with norm

$$\| x \| = \left(\int | x(t) |^p \, dt\right)^{1/p}, \qquad 1 \leqq p < \infty;$$

and the Banach space (M) of almost everywhere bounded measurable functions $x(t)$, with norm

$$\| x \| = \text{true max } | x(t) |, \qquad p = \infty.$$

1. **Projections and involutions.** A *projection* in a Banach space B is any linear operation P in the space which is such that $P^2 = P$. (In particular, P may be the identity or the zero operation.) The range of a projection P is the closed linear subspace consisting of all elements $x \in B$ such that $Px = x$. In order that two linear operations P and P' in a given Banach space be projections on the same subspace, it is obviously necessary and sufficient[5] that $P'P = P$, $PP' = P'$.

An *involution* in a Banach space B is any linear operation U in the space

[4] By "function" we mean here an equivalence class of functions, two functions belonging to the same class if and only if they differ only on a set of zero measure.

[5] We verify the sufficiency as follows. Let l_P and $l_{P'}$ denote the ranges of P and of P'. Then $P'P = P$ implies $l_{P'} \supset l_P$, $PP' = P'$ implies $l_P \supset l_{P'}$; and for any $x \in l_P = l_{P'}$, $Px = P'x = x$. (Or $P'PP' = PP' = P'$, $(P')^2 = P'$; similarly $P^2 = P$.)

which is such that $U^2 = I$ (where I is the identity operation). The subspace of an involution is the closed linear subspace of all elements $x \in B$ which are such that $Ux = x$. We observe first the close connection between projections and involutions.

THEOREM 1.1. *A n.s.c. (necessary and sufficient condition) that there exist a projection of a space B on a subspace B_1 is that there exist an involution U in B having B_1 as its subspace. To any projection P corresponds an involution U given by the equation $U = 2P - I$; and to any involution U corresponds a projection P given by the equation $P = \frac{1}{2}(U + I)$; the subspaces of P and of U are identical.*

Proof. Suppose P exists, and let $U = 2P - I$. Since $P^2 - P = P(P - I) = 0$, we obtain $[\frac{1}{2}(U + I)][\frac{1}{2}(U - I)] = 0$, or $U^2 = I$. Conversely, if U exists and P is defined by $P = \frac{1}{2}(U + I)$, then $P^2 = P$. Since $U = 2P - I$, if P is bounded, U is bounded, and $| U | \leqq 2 | P | + 1$. If U is bounded, P is bounded, and $| P | = \frac{1}{2} | U + I | \leqq \frac{1}{2}(| U | + 1)$.

THEOREM 1.2. *Let V be a linear operation. If $P = \frac{1}{2}(I + U)$ is a projection on a subspace B_1, then a n.s.c. for $P' = \frac{1}{2}(I + U + V)$ to be a projection on the same subspace B_1 is that*

$$UV = -VU = V.$$

Remark. In the finite-dimensional case, where linear transformations are represented by matrices, the condition of Theorem 1.2 implies that the trace of the matrix $V = UV = -VU$ is zero.

Proof. Suppose P and P' are projections on the same subspace. Then $PP' = P'$ and $P'P = P$; or $\frac{1}{4}(I + U)(I + U + V) = \frac{1}{2}(I + U) + \frac{1}{4}(V + UV) = \frac{1}{2}(I + U + V)$, or $UV = V$; in the same way $V = -VU$; and the condition is necessary.

Conversely, if $UV = -VU = V$, then since P is a projection, $\frac{1}{4}(I + U)(I + U + V) = \frac{1}{2}(I + U) + \frac{1}{4}(2V)$, or $PP' = P'$; $\frac{1}{4}(I + U + V)(I + U) = \frac{1}{2}(I + U)$, or $P'P = P$; and the condition is sufficient.

We shall find it more convenient to deal with involutions than with projections. Accordingly we introduce certain quantities for involutions which are analogous to quantities which Murray defined for projections. (See [6], p. 140.)

DEFINITION 1.1. Let l be a subspace of an n-dimensional Minkowski space l_n. By $K(l)$ we denote the minimum of the norms of all possible involutions of l_n in l. An involution U whose norm $| U | = K(l)$ is called a *minimal* involution. We denote by $\bar{K}(l_n)$ the maximum of $K(l)$ over all subspaces $l \subset l_n$. Corresponding K's are defined for closed linear subspaces of general Banach spaces by replacing "maximum" and "minimum" respectively by "l.u.b." and "g.l.b.", with the understanding that K and $\bar{K}$ may take on the value $+\infty$. (By $K(B_1) = +\infty$ we mean that no involution exists in the subspace B_1.)

We retain Murray's C and $\bar{C}$ notation for the precisely similar quantities for projections. By Theorem 1.1, if $K(B_1) = \infty$, then $C(B_1) = \infty$; i.e., if no

involution in the subspace B_1 exists, then no projection on B_1 exists; and conversely.

LEMMA 1.1. *For any Minkowski space l_n, if l is a subspace,*

$$\tfrac{1}{2}(K(l) - 1) \leqq C(l) \leqq \tfrac{1}{2}(K(l) + 1);$$

and

$$\tfrac{1}{2}(\bar{K}(l_n) - 1) \leqq \bar{C}(l_n) \leqq \tfrac{1}{2}(\bar{K}(l_n) + 1).$$

This lemma is a consequence of Theorem 1.1.

LEMMA 1.2. *Given any closed linear subspace l of a Banach (Minkowski) space L. Then $\bar{C}(L) \geqq \bar{C}(l)$; and if l_1 is any closed linear subspace of l, $C(l_1)$ in $L \geqq C(l_1)$ in l.*

This lemma is obvious, since l_1, being closed in l, is also closed in L; and if a projection of L on l_1 exists, it is in particular a projection of l on l_1.

LEMMA 1.3. *If l_n and $\mathbf{l}_n$ are conjugate Minkowski spaces,*

$$\bar{C}(\mathbf{l}_n) \leqq \bar{C}(l_n) + 1 \quad \textit{and} \quad \bar{K}(l_n) = \bar{K}(\mathbf{l}_n).$$

Proof. Let f be a functional on l_n and let (f, x) represent the value assumed by f for the element x. Then $x \in l_n$, $f \in \mathbf{l}_n$. Suppose l is a subspace of l_n. Then the set of all $f \in \mathbf{l}_n$ such that $(f, x) = 0$ for all $x \in l$ is a subspace $l(\perp)$ of $\mathbf{l}_n$. If U is any involution in l, U^* the adjoint operation, then $-U^*$ is an involution in $l(\perp)$, and $| U | = | U^* |$. Therefore $K(l) = K[l(\perp)]$, and $\bar{K}(l_n) = \bar{K}(\mathbf{l}_n)$. That $\bar{C}(\mathbf{l}_n) \leqq \bar{C}(l_n) + 1$ now follows from Lemma 1.1.

2. **Projections in the spaces l_p and L_p.** Although considerations similar to those introduced in this section will be used later for more general spaces, so that possibly this section could be incorporated into a later section, we consider the l_p-spaces separately to show how easily our methods yield the results of Murray. We believe also that our methods afford an improved insight into the situation as regards projections in l_p.

Following Murray, we approach the spaces l_p by considering the n-dimensional spaces $l_{p,n}$. Murray showed that the existence of a subspace $l \subset l_p$ such that $C(l) = \infty$ is implied by $\lim_{n\to\infty} \bar{C}(l_{p,n}) = \infty$ (or $\lim_{i\to\infty} \bar{C}(l_{p,n_i}) = \infty$ for any subsequence $\{n_i\}$ of $\{n\}$). His proof that the limit is ∞ used the sequence of spaces $\{l_{p,3^\nu}\}$ $(\nu = 1, 2, 3, \cdots)$, together with a certain sequence of subspaces $\{l^\nu\}$, $l^\nu \subset l_{p,3^\nu}$ for each ν, the dimension of each l^ν being $k_\nu = 2^\nu$. For these subspaces, however, $K(l^\nu)$ does not have the maximum possible value [i.e., $K(l^\nu) < \bar{K}(l_{p,3^\nu})$]. Our proof uses similarly the sequence of spaces $\{l_{p,2^\nu}\}$, together with a system of subspaces $\{l^\nu\}$ of a different nature, for which the dimension k_ν is $2^{\nu-1} = \tfrac{1}{2}n$, and for which $K(l^\nu)$ does have its maximum value. Note that for Murray's subspaces, the ratio of the dimension of l^ν to that of the corresponding space $l_{p,n}$ is $(\tfrac{2}{3})^\nu$, which approaches 0 as $\nu \to \infty$, while for our subspaces the ratio is constant and equal to $\tfrac{1}{2}$. By our more appropriate choice

of subspaces, we are able not only to show much more briefly that the limit is ∞, but also to discuss the precise rate of growth of $\bar{C}(l_{p,n})$.

THEOREM 2.1. *In any space* $l_{p,n}$ $(1 \leqq p \leqq \infty)$ *of dimension* $n = 2^\nu$, ν *an integer, there is a subspace* l *such that*

$$\bar{K}(l_{p,n}) \geqq K(l) \geqq n^{|1/p - \frac{1}{2}|}; \qquad \bar{C}(l_{p,n}) \geqq C(l) \geqq \tfrac{1}{2}(n^{|1/p-\frac{1}{2}|} - 1).$$

Proof. Consider the set of matrices $\{\beta_\nu\}$ defined inductively as follows:

$$\beta_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}; \qquad \beta_\nu = \beta_{\nu-1} \quad \text{with } \pm\beta_1 \text{ substituted for } \pm 1.$$

For any $n = 2^\nu$, the matrix $U = n^{-\frac{1}{2}}\beta_\nu$ is a symmetric orthogonal matrix, i.e., an (orthogonal) involutoric matrix.[6] Therefore if $x = (x_1, \cdots, x_n)$, $x \in l_{p,n}$, the linear transformation[7] defined by $y = Ux$ is an involution in $l_{p,n}$.

As noted by Murray, by conjugate spaces it is sufficient to consider only the case of $p < 2$. (Lemma 1.3.) A direct proof for the case of $p > 2$ is as easy as the proof which follows for $p < 2$.

Let l be the subspace of U, where $U = \{u_{ij}\} = n^{-\frac{1}{2}}\beta_\nu$ as above. Let $Y_1 = (1, 0, \cdots, 0)$, $Y_2 = (0, 1, 0, \cdots, 0), \cdots, Y_n = (0, 0, \cdots, 0, 1)$. We now make use of the remark following Theorem 1.2. If $U + V$ is any involution in l, where $V = \{v_{ij}\}$, we must have trace $(V) = 0$, and therefore for at least one k, $v_{kk} \geqq 0$. Since $V = UV$ and U is symmetric, $v_{kk} = \sum_i u_{ki}v_{ik} = \sum_i u_{ik}v_{ik}$. Also since $U^2 = I$, $1 = \sum_i u_{ki}u_{ik} = \sum_i u_{ik}u_{ik}$. By Hölder's inequality, for any k such that $v_{kk} \geqq 0$, we have

$$1 \leqq 1 + v_{kk} = \sum_i u_{ik}\cdot(u_{ik} + v_{ik}) \leqq n^{\frac{1}{2}-1/p} \,||\, \{u_{ik} + v_{ik}\} \,||$$
$$= n^{\frac{1}{2}-1/p} \,||\, (U + V)\, Y_k \,|| \leqq n^{\frac{1}{2}-1/p}\cdot |\, U + V \,|,$$

or $|\, U + V \,| \geqq n^{1/p - \frac{1}{2}}$. Thus $K(l) \geqq n^{1/p-\frac{1}{2}}$. The second statement of Theorem 2.1 follows by Lemma 1.1.

The existence of subspaces of l_p for which there are no projections may now

[6] If α and β are any two symmetric or orthogonal (unnormalized) matrices of ± 1's, then the matrix obtained by substituting $\pm\beta$ for ± 1 in α is respectively a symmetric or orthogonal matrix of ± 1's; its order is the product of the orders of α and β. The verification is immediate. Thus since β_1 is symmetric and orthogonal (unnormalized), β_ν is likewise; and with the normalizing factor $n^{-\frac{1}{2}}$, $U = n^{-\frac{1}{2}}\beta_\nu$ is an orthogonal matrix, so that $y = Ux$ is an orthogonal transformation in $l_{2,n}$.

A n.s.c. for orthogonality of a matrix U is that $U' = U^{-1}$, where U' is the transposed matrix (the transposed conjugate matrix in the complex case). Therefore if U is orthogonal, U symmetric (Hermitian) implies $U = U^{-1}$, or $U^2 = I$; conversely U orthogonal and involutoric implies that U is symmetric (Hermitian).

[7] In any expression of the form $y = Ux$ we think of $x = (x_1, \cdots, x_n)$ as a "column matrix", i.e., a matrix of n rows and one column (although for convenience the x_i's are written in a horizontal row). Then $y = (y_1, \cdots, y_n)$ is the column matrix obtained by multiplying the column matrix x from the left by the matrix U, according to the usual rule for matrix multiplication.

be seen as follows. We regard l_p as the sum of a sequence of subspaces $\{l_{p,2^\nu}\}$, where $\nu \to \infty$, and the $l_{p,2^\nu}$ do not intersect, except in the 0 of l_p. We define a closed linear subspace l of l_p as the sum of the sequence of subspaces $\{l^\nu\}$ given by Theorem 2.1. (This process is described in detail later, in a more general situation; namely, in the proof of Theorem 4.2.) That there exists no projection for the subspace l is implied immediately by Theorem 2.1 and the following lemma of Murray.

LEMMA 2.1. (Murray.) *Given a Banach space L with a closed linear subspace M on which there exists a projection P of norm 1; let N be the corresponding complementary subspace to M. Let l be any closed linear subspace of L which is such that if $x \in l$, then $x = y + z$, where $y \in M \cdot l$, $z \in N \cdot l$. Let l_1 be the closed linear subspace $M \cdot l$. Then $C(l_1) \leqq C(l)$.*

For the proof of this lemma, the reader is referred to p. 140 of [6].

As pointed out by Murray, the space l_p is equivalent to a closed linear subspace of the function space L_p, so that the existence of a subspace $l \subset l_p$ such that $C(l) = \infty$ implies the same for the space L_p. (Lemma 1.2.) That $\lim_{n\to\infty} \bar{C}(l_{p,n}) = \infty$ follows from Theorem 2.1, since by Lemma 1.2 $\bar{C}(l_{p,n})$ is monotone.

The discussion above has included the cases $p = 1$ and $p = \infty$.[8]

3. A closer estimate of $\bar{C}(l_{p,n})$. The affine ratio ρ.

In Euclidean space or in Hilbert space, a n.s.c. that a projection P be an *orthogonal* projection is that $| P | = 1$ (excluding $P \equiv 0$); and such a projection exists on every closed linear subspace. A similar statement is true for involutions. The projection P corresponding according to Theorem 1.1 to an orthogonal involution U is an orthogonal projection, and conversely.

[8] In the case $p = \infty$, the space $l_\infty = (m)$ is equivalent to the closed linear subspace of $L_\infty = (M)$ determined by the functions $y_i(t) = 1$ for $2^{-i} \leqq t \leqq 2^{-i+1}$, otherwise $y_i(t) = 0$ $(i = 1, 2, \cdots)$. Just as in the case of finite p, Theorem 2.1, Lemma 2.1, and Lemma 1.2 imply the existence of subspaces $l' \subset (m)$ such that $C(l') = \bar{C}[(m)] = \infty$, and $l'' \cong l'$, $l'' \subset (M)$, such that $C(l'') = \bar{C}[(M)] = \infty$.

The spaces (c_0) of sequences convergent to 0, and (c) of convergent sequences, are subspaces of the space $l_\infty = (m)$. Both (c) and (c_0) contain all of the spaces $l_{\infty,n}$ as subspaces, and exactly the same considerations as used above for the case of l_p-spaces yield the results that there exist closed linear subspaces $l_0 \subset (c_0)$ and $l \subset (c)$, such that $C(l_0) = \bar{C}[(c_0)] = \infty$ and $C(l) = \bar{C}[(c)] = \infty$. Moreover, in (m), $l_0 \subset l \subset l'$, and l_0 and l are additional closed linear subspaces for which $C(l_0) = C(l) = \infty$.

Finally, we remark that the spaces (C) of continuous functions on [0, 1], and $(C)^p$ of functions having a continuous p-th derivative on [0, 1], likewise contain closed linear subspaces for which there are no projections. This result follows for (C) by the preceding paragraph and Lemma 1.2, since (C) obviously contains a closed linear subspace equivalent to (c). (The result for (C) is implied also by our previous results and Lemma 1.2, in view of Theorem 9, p. 185 of [2].) The result follows for $(C)^p$ from that for (C), since according to Theorem 7, p. 184 of [2], $(C)^p$ and (C) are isomorphic.

The above remarks dispose of all of the blanks in row (7) of the table on page 245 of [2].

For a given Banach space B, it may be possible to introduce a Hilbert norm $\| x \|_2$ in such a way that, for all $x \in B$, the inequalities

$$\| x \|_2 \leqq \| x \| \leqq C \cdot \| x \|_2$$

are satisfied, where $\| x \|$ is the given norm. (This is always possible when the Banach space is a Minkowski space.) In such a case we may consider orthogonal projections with respect to the Hilbert norm; and they are projections with respect to the given norm. By the norm $| P |$ of an orthogonal projection P we shall mean the norm of the corresponding projection in the given Banach (Minkowski) space. Similarly for orthogonal involutions.

THEOREM 3.1. *Given a Banach (Minkowski) space which, in addition to the given norm, admits a Hilbert norm as above. Then, if* $P = \frac{1}{2}(U + I)$ *is any orthogonal projection in the space,*

$$| P | \leqq \tfrac{1}{2} + \tfrac{1}{2}C \qquad \text{and} \qquad | U | \leqq C.$$

Proof. For all x

$$\| Ux \|^2 \leqq C^2 \cdot \| Ux \|_2^2 = C^2 \cdot (Ux, Ux) = C^2 \cdot \| x \|_2^2 \leqq C^2 \cdot \| x \|^2,$$

or $| U | \leqq C$; and

$$| P | = \tfrac{1}{2} | U + I | \leqq \tfrac{1}{2}(| U | + 1) \leqq \tfrac{1}{2} + \tfrac{1}{2}C.$$

DEFINITION 3.1. Given any Banach space B. If it is possible to introduce a Hilbert norm as above, we define ρ as the g.l.b. of C over all possible ways of introducing a Hilbert norm. Otherwise we take $\rho = \infty$. The quantity ρ will be called the *affine ratio* of the space B.

If ρ is finite, the Banach space B is isomorphic to either Euclidean, ordinary Hilbert, or a hyper-Hilbert space; and conversely. If the space B is isomorphic to ordinary Hilbert (or Euclidean) space, it may be shown[9] that the g.l.b. of

[9] The verification of the statement is as follows. For a Hilbert norm $\| x \|_n$, let C_n be the smallest constant such that $\| x \|_n \leqq \| x \| \leqq C_n \cdot \| x \|$ for all x. Choose a sequence of Hilbert norms $\| x \|_n$ such that $\lim_{n \to \infty} C_n = \rho$. By separability we have a countable dense subset $\{x_k\}$ in B; for each k and n $\| x_k \|_n \leqq \| x_k \| \leqq C_n \cdot \| x_k \|_n$. For a fixed k, and any subsequence $\{n_i\}$ of $\{n\}$, the $\| x_k \|_{n_i}$ are a bounded sequence and therefore they contain a convergent subsequence. Hence by a diagonal procedure we may extract a subsequence $\{n_i\}$ of $\{n\}$ such that for every k, $\lim_{i \to \infty} \| x_k \|_{n_i}$ exists; let this limit be denoted by $\| x_k \|_2$. Change notation so that the sequence $\{n_i\}$ becomes $\{n\}$. Then for any $x \in B$,

$$| \| x \|_n - \| x \|_{n'} | \leqq \| x - x_k \|_n + | \| x_k \|_n - \| x_k \|_{n'} | + \| x_k - x \|_{n'}$$
$$\leqq 2 \| x - x_k \| + | \| x_k \|_n - \| x_k \|_{n'} |,$$

so that $\{\| x \|_n\}$ is a Cauchy sequence and therefore a convergent sequence; let its limit be denoted by $\| x \|_2$. Obviously $\| x \|_2$ is a norm defined on B, such that $\| x \|_2 \leqq \| x \| \leqq \rho \| x \|_2$. The n.s.c. for a Hilbert norm of von Neumann and Jordan [5] is that

$$\| x - y \|^2 + \| x + y \|^2 = 2(\| x \|^2 + \| y \|^2).$$

This condition is satisfied by $\| x \|_n$ for all n; therefore it is satisfied by $\| x \|_2$, and $\| x \|_2$ is a Hilbert norm.

Definition 3.1 is actually attained, i.e., that it is possible to introduce a Hilbert norm $\| x \|_2$ such that for all $x \in B$

$$\| x \|_2 \leqq \| x \| \leqq \rho \cdot \| x \|_2 .$$

It may easily be shown also that if ρ (finite or ∞) is the affine ratio for a Banach space B, and if $\bar{\rho}$ is the affine ratio for the conjugate space $\bar{B}$, then $\bar{\rho} = \rho$.

COROLLARY TO THEOREM 3.1. *For any Banach (Minkowski) space L,*

$$\bar{C}(L) \leqq \tfrac{1}{2} + \tfrac{1}{2}\rho \quad \textit{and} \quad \bar{K}(L) \leqq \rho.$$

This corollary will enable us to verify that the lower bounds of Theorem 2.1 are essentially the best possible estimates of the quantities $\bar{C}(l_{p,n})$ and $\bar{K}(l_{p,n})$.

THEOREM 3.2. *If $n = 2^r$, $\bar{K}(l_{p,n}) = n^{|1/p - \frac{1}{2}|}$, and*

$$\tfrac{1}{2}n^{|1/p-\frac{1}{2}|} - \tfrac{1}{2} \leqq \bar{C}(l_{p,n}) \leqq \tfrac{1}{2}n^{|1/p-\frac{1}{2}|} + \tfrac{1}{2} \qquad (1 \leqq p \leqq \infty)\cdot$$

Proof. In any space $l_{p,n}$, let us introduce a Hilbert norm by

$$\| x \|_2 = \Big(\sum_{i=1}^{n} | x_i |^2\Big)^{\frac{1}{2}}, \qquad \text{where } x = (x_1, \cdots, x_n),\ x \in l_{p,n} .$$

Consider first the case $p < 2$. Then we have the inequality $(\sum | x_i |^2)^{\frac{1}{2}} \leqq (\sum | x_i |^p)^{1/p}$.[10] Also, applying Hölder's inequality

$$\sum a_i b_i \leqq \Big(\sum | a_i |^q\Big)^{1/q} \cdot \Big(\sum | b_i |^{q'}\Big)^{1/q'} \qquad \left(\frac{1}{q} + \frac{1}{q'} = 1\right),$$

with $q' = 2/p$, we obtain

$$\Big(\sum | x_i |^p\Big) = \sum (1 \cdot | x_i |^p) \leqq n^{1-\frac{1}{2}p} \cdot \Big(\sum | x_i |^{p \cdot 2/p}\Big)^{\frac{1}{2}p}$$

or

$$\Big(\sum | x_i |^p\Big)^{1/p} \leqq n^{1/p - \frac{1}{2}} \cdot \Big(\sum | x_i |^2\Big)^{\frac{1}{2}}.$$

Thus $\| x \|_2 \leqq \| x \| \leqq C \cdot \| x \|_2$, where $n^{|1/p-\frac{1}{2}|} = C \geqq \rho$. Therefore by Theorem 2.1 and the above corollary, we have

$$\rho = n^{|1/p - \frac{1}{2}|} = \bar{K}(l_{p,n}).$$

For $q > 2$, by Lemma 1.3 $\bar{K}(l_{p,n}) = \bar{K}(l_{q,n})$, where $p^{-1} + q^{-1} = 1$. Thus if $\bar{\rho}$ is the affine ratio for $l_{q,n}$, $\bar{\rho} = \rho = n^{1/p - \frac{1}{2}} = n^{|1/q - \frac{1}{2}|}$. Therefore for any p in the range $1 \leqq p \leqq \infty$, we have $\rho = n^{|1/p - \frac{1}{2}|} = \bar{K}(l_{p,n})$. Lemma 1.1 now implies the statement concerning $\bar{C}(l_{p,n})$.

THEOREM 3.21. *In any space $l_{p,n}$ of dimension n such that there exists a symmetric orthogonal (unnormalized) matrix of ± 1's of order n, the statements of*

[10] For this inequality see e.g. Bohnenblust, *Functions of Real Variables* (Princeton Notes), Ann Arbor, Mich., 1937, p. 19.

Theorem 3.2 *are true. In particular, Theorem* 3.2 *is true if n is* 12, *or* $12 \cdot 2^\nu$, *ν an integer.*

Proof. The proof of Theorem 2.1, on which Theorem 3.2 depends, obviously requires only the fact that β_ν in $U = n^{-\frac{1}{2}}\beta_\nu$ is a symmetric orthogonal (unnormalized) matrix of ± 1's. Therefore to prove Theorem 3.21, by footnote 6 we need only to exhibit a symmetric orthogonal (unnormalized) matrix of ± 1's of order 12. The following is such a matrix, where $+$ and $-$ represent respectively $+1$ and -1.

$$
\begin{array}{cccccccccccc}
+ & + & + & + & + & + & + & + & + & + & + & + \\
+ & + & + & + & + & + & - & - & - & - & - & - \\
+ & + & + & - & - & - & + & + & + & - & - & - \\
+ & + & - & - & + & - & - & + & - & + & + & - \\
+ & + & - & + & - & - & + & - & - & + & - & + \\
+ & + & - & - & - & + & - & - & + & - & + & + \\
+ & - & + & - & + & - & - & - & + & + & - & + \\
+ & - & + & + & - & - & - & + & - & - & + & + \\
+ & - & + & - & - & + & + & - & - & + & + & - \\
+ & - & - & + & + & - & + & - & + & - & + & - \\
+ & - & - & + & - & + & - & + & + & + & - & - \\
+ & - & - & - & + & + & + & + & - & - & - & +
\end{array}
$$

Paley has shown[11] that an orthogonal (unnormalized) matrix of ± 1's, or even three orthogonal rows of ± 1's, is possible only if the order n is divisible by 4 (except $n = 1$ and 2). He conjectures that such a matrix is in fact possible for any order n which is a multiple of 4. The above matrix shows that the additional requirement of symmetry is not sufficient to force n to be a power of 2, as might have been expected. It may be that a symmetric matrix likewise is possible whenever n is a multiple of 4.

In consequence of Theorem 3.2, we have the following asymptotic relationship:

$$2\bar{C}(l_{p,n}) \sim \bar{K}(l_{p,n}) \sim n^{|1/p - \frac{1}{2}|}.$$

The following two theorems, in which are obtained a still more precise estimate of $\bar{C}(l_{p,n})$, and the exact values of $\bar{C}(l_{1,n})$ and $\bar{C}(l_{\infty,n})$, are of interest.

THEOREM 3.3. *In any space $l_{p,n}$ $(1 < p < \infty,\ p \neq 2)$ of dimension $n = 4^\nu$, there is a subspace l which is such that*

$$\tfrac{1}{2} n^{|1/p - \frac{1}{2}|} < C(l) \leqq \bar{C}(l_{p,n}) < \tfrac{1}{2} + \tfrac{1}{2} n^{|1/p - \frac{1}{2}|}.$$

[11] R. E. A. C. Paley, *On orthogonal matrices*, Journal of Math. and Phys., Mass. Inst. Tech., vol. 12(1933), pp. 311–312.

Proof. Let us define a set of matrices $\{\alpha_\nu\}$ as follows:

$$\alpha_1 = -\begin{vmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{vmatrix};$$

for each successive ν, α_ν is defined by substituting $\pm\alpha_1$ for ± 1 in $\alpha_{\nu-1}$.[12] Then for any $n = 4^\nu$, the matrix $U = n^{-\frac{1}{2}}\alpha_\nu = \{u_{ij}\}$ is involutoric,[13] and $P = \frac{1}{2}(I + U)$ is a projection on the subspace l of the involution U.

Consider first the case $p < 2$. Define $\{Y_i\}$ as in the proof of Theorem 2.1. Let $\sum_i'$ denote summation over i omitting j. Then for every j, by the orthogonality of U,

$$\| PY_j \| = \| \tfrac{1}{2}(I + U)Y_j \| = \tfrac{1}{2}[(1 + u_{jj})^p + \textstyle\sum_i' | u_{ij} |^p]^{1/p}$$

$$= \tfrac{1}{2}\left[(1 + n^{-\frac{1}{2}})^p + \frac{n-1}{n^{\frac{1}{2}p}}\right]^{1/p} > \tfrac{1}{2}[n \cdot n^{-\frac{1}{2}p}]^{1/p} = \tfrac{1}{2}n^{1/p-\frac{1}{2}}.$$

We show by contradiction that $C(l) > \frac{1}{2}n^{1/p-\frac{1}{2}}$. For suppose that $C(l) \leqq \frac{1}{2}n^{1/p-\frac{1}{2}}$. Then by hypothesis there exists a projection $P' = \frac{1}{2}(I + U + V)$ such that $\| P'Y_j \| < \| PY_j \|$ for all j. We may suppose furthermore that the v_{ij}'s in $V = \{v_{ij}\}$ are as small as desired, since if P'' is any projection such that $\| P''Y_j \| < \| PY_j \|$ for all j, the projection $P' = P_t = tP'' + (1 - t)P$ has the same property for any t in the range $0 < t \leqq 1$.

Writing out $\| P'Y_j \|^p < \| PY_j \|^p$, we have

$$|1 + u_{jj}|^p + \textstyle\sum_i' | u_{ij} |^p > | 1 + u_{jj} + v_{jj} |^p + \textstyle\sum_i' | u_{ij} + v_{ij} |^p.$$

Also

$$| u_{ij} + v_{ij} |^p = | n^{-\frac{1}{2}}\cdot\text{sign } u_{ij} + v_{ij}\cdot n^{-\frac{1}{2}}\cdot n^{-\frac{1}{2}}\cdot n |^p = n^{-\frac{1}{2}p} | 1 + v_{ij}u_{ij}n |^p.$$

By use of the mean value theorem, it may be easily proved that $(a + x)^p > a^p + pa^{p-1}x$, for $a > 0$, $| x | < a$ if $x < 0$, $p > 1$. Applying this inequality, we have

$$(1 + n^{-\frac{1}{2}})^p + \frac{n-1}{n^{\frac{1}{2}p}} > | 1 + n^{-\frac{1}{2}} + v_{jj} |^p + n^{-\frac{1}{2}p} \textstyle\sum_i' | 1 + v_{ij}\cdot u_{ij}\cdot n |^p$$

$$> (1 + n^{-\frac{1}{2}})^p + p\cdot v_{jj}\cdot(1 + n^{-\frac{1}{2}})^{p-1} + \frac{n-1}{n^{\frac{1}{2}p}} + pn^{1-\frac{1}{2}p} \textstyle\sum_i' v_{ij}u_{ij},$$

[12] The matrices $\{\alpha_\nu\}$ and $\{\beta_\nu\}$ have been previously considered. See e.g. J. E. Littlewood, *On bounded bilinear forms*, Quarterly Journal of Mathematics (Oxford series), vol. 1(1930), p. 172, and the reference to Paley already cited.

[13] See footnote 6.

or

$$0 > p \cdot v_{jj}(1 + n^{-\frac{1}{2}})^{p-1} - pn^{1-\frac{1}{2}p} v_{jj} u_{jj} + pn^{1-\frac{1}{2}p} \sum_i v_{ij} u_{ij}.$$

Summing over j and remembering that U is symmetric and that $u_{jj} = n^{-\frac{1}{2}}$ for all j, we obtain

$$0 > p[(1 + n^{-\frac{1}{2}})^{p-1} - n^{\frac{1}{2}-\frac{1}{2}p}] \text{ trace } (V) + pn^{1-\frac{1}{2}p} \text{ trace } (UV) = 0,$$

by Theorem 1.2. This is the desired contradiction.

We consider now the case $p > 2$. For each i, we define $X_i = (\text{sign } u_{i1}, \cdots, \text{sign } u_{in})$. Then $\| X_i \| = n^{1/p}$. By the orthogonality of U,

$$\begin{aligned} \| PX_j \| = \tfrac{1}{2} \| (U + I)X_j \| &= \tfrac{1}{2}[\textstyle\sum_i' | \text{sign } u_{ji} |^p + (n^{\frac{1}{2}} + 1)^p]^{1/p} \\ &= \tfrac{1}{2}[(n - 1) + (n^{\frac{1}{2}} + 1)^p]^{1/p} > \tfrac{1}{2}[n^{\frac{1}{2}p} + n]^{1/p} > \tfrac{1}{2}n^{\frac{1}{2}}, \end{aligned}$$

or $\| PX_j \| > \frac{1}{2}n^{\frac{1}{2}-1/p} \cdot \| X_j \|$. Thus $| P | > \frac{1}{2}n^{\frac{1}{2}-1/p}$.

We show by contradiction that $C(l) > \frac{1}{2}n^{\frac{1}{2}-1/p}$. For suppose that $C(l) \leqq \frac{1}{2}n^{\frac{1}{2}-1/p}$. Then, from the previous paragraph, by hypothesis there is a projection $P' = \frac{1}{2}(I + U + V)$ such that, for each j, $\| P'X_j \|^p < \| PX_j \|^p$. Writing this out, using the above equality for $\| PX_j \|$ and the previous inequality derived from the mean value theorem, we have

$$\begin{aligned} (n - 1) + (n^{\frac{1}{2}} + 1)^p &> \| (I + U + V)X_j \|^p \\ &= | n^{\frac{1}{2}} + \textstyle\sum_i v_{ji} \text{ sign } u_{ji} + 1 |^p + \sum_k' | \text{sign } u_{jk} + \sum_i v_{ki} \text{ sign } u_{ji} |^p \\ &> (n^{\frac{1}{2}} + 1)^p + p(n^{\frac{1}{2}} + 1)^{p-1} \cdot \textstyle\sum_i v_{ji} \text{ sign } u_{ji} + (n - 1) \\ &\qquad + \textstyle\sum_k' (p \text{ sign } u_{jk} \cdot \sum_i v_{ki} \text{ sign } u_{ji}), \end{aligned}$$

where $\sum_k'$ denotes summation over k omitting j. Transposing $(n^{\frac{1}{2}} + 1)^p + (n - 1)$ from the last to the first member of these inequalities, summing over j, and using the symmetry of U, we have

$$0 > pn^{\frac{1}{2}}(n^{\frac{1}{2}} + 1)^{p-1} \text{ trace } (VU) + p \sum_j \sum_k' \sum_i v_{ki} \text{ sign } u_{jk} \text{ sign } u_{ji}.$$

Now since

$$\sum_j \sum_i v_{ji} \text{ sign } u_{ji} = n^{\frac{1}{2}} \text{ trace } (VU) = \sum_j \text{sign } u_{jj} \cdot \sum_i v_{ji} \text{ sign } u_{ji} = 0$$

by Theorem 1.2, we may allow the k-summation in this inequality to include $k = j$, and by the orthogonality of U we then obtain

$$\begin{aligned} 0 > \sum_j \sum_k \sum_i v_{ki} \text{ sign } u_{jk} \text{ sign } u_{ji} &= \sum_i \sum_j v_{ii} \text{ sign } u_{ji} \text{ sign } u_{ji} \\ &= \sum_i nv_{ii} = n \text{ trace } (V) = 0 \end{aligned}$$

by Theorem 1.2. This is the desired contradiction.

THEOREM 3.4. *In any space $l_{1,n}$ or $l_{\infty,n}$ of dimension $n = 4^\nu$, ν an integer, there is a subspace l which is such that*

$$C(l) = \bar{C}(l_{1,n}) = \bar{C}(l_{\infty,n}) = \tfrac{1}{2} + \tfrac{1}{2}n^{\frac{1}{2}} = \tfrac{1}{2} + \tfrac{1}{2}\rho.$$

The dimension of the subspace is $\frac{1}{2}n + \frac{1}{2}n^{\frac{1}{2}} = \frac{1}{2}4^\nu + \frac{1}{2}2^\nu$.

The proof of Theorem 3.4 is omitted, since the main part of the proof is similar to the proof of Theorem 3.3. The dimension of the subspace l is the trace of the matrix $\frac{1}{2}(n^{-\frac{1}{2}}\alpha_\nu + I)$. Similarly, the dimension of the subspace of Theorem 2.1 is $\frac{1}{2}n$.

4. **Symmetric sequence spaces.** In the remainder of this paper we study Banach spaces of sequences more general than the spaces l_p.

An infinite sequence $x = \{x_i\}$ in which all but a finite number of the coordinates x_i are zero will be called a "finite" sequence. We shall use the briefer notation $x = (x_1, \cdots, x_n)$ for the "finite" sequence $x = (x_1, \cdots, x_n, 0, 0, \cdots)$. For any infinite sequence $\{x_i\}$, let $\{x_i\}_n$ denote $(x_1, \cdots, x_n, 0, 0, \cdots) = (x_1, \cdots, x_n)$. The set of all "finite" sequences, with the usual rules of operation for sequences, is obviously a linear space; let this linear space be denoted by S_f.

Suppose that we are given any norm $\| x \|'$ on S_f. Define a second norm by $\| x \| = \max_{n \geq 1} \| \{x_i\}_n \|'$ for each $x \in S_f$. A Banach space then arises in the following manner: We add to the normed linear space S_f all infinite sequences $x = \{x_i\}$ such that $\text{l.u.b.}_{1 \leq n < \infty} \| \{x_i\}_n \|$ is finite, and we define $\| x \|$ to be the value of this expression. By the properties of the norm on S_f, the set of finite and infinite sequences so obtained, with norm $\| x \|$, is a normed linear space S. A proof that S is complete, and so a Banach space, may be obtained by repetition in greater generality of a classical procedure (that used, for example, to show the completeness of the sequence spaces l_p).[14] This general completeness proof is omitted here, since the procedure is so well known.

Remark. If we are given any two norms $\| x \|_1'$ and $\| x \|_2'$ on S_f which are isomorphic, with constants C_1 and C_2, then obviously the corresponding Banach spaces contain the same infinite sequences, and on the common linear space the Banach norms are isomorphic, with the same C_1 and C_2.

[14] For this procedure, see e.g. Bohnenblust, op. cit., pp. 96–97.

Because of the redefinition of the norm in S_f, obviously $\| x \| = \lim_n \| \{x_i\}_n \|$. If we define $\| x \| = \text{l.u.b.} \| \{x_i\}_n \|'$ only for infinite sequences $x \in S$, and $\| x \| = \| x \|'$ for all $x \in S_f$, then the triangle property of the norm is not necessarily satisfied by $\| x \|$ on the extended space S. If S' is any Banach space of infinite dimension, and if $\{X_n\}$ is any sequence of linearly independent elements of S', then the linear subspace of all finite combinations $\sum x_i X_i$ is a space S_f, and the closure $\bar{S}_f$ of S_f in S' is equivalent to a separable Banach space of sequences. But for $x = \{x_i\} \in \bar{S}_f$, it is not necessarily true either that $\text{l.u.b.} \| \{x_i\}_n \|$ is finite, or that $\lim_n \| \{x_i\}_n \|$ exists and > 0.

THEOREM 4.1. *If two Banach spaces of sequences, which correspond to two norms on S_f in the way described above, contain the same sequences, they are isomorphic. The isomorphism is given by the identity transformation in the common linear space of sequences.*

Proof. Let $\| x \|_1$ and $\| x \|_2$ be the two Banach norms on the common linear space of sequences. Define a third norm $\| x \| = \| x \|_1 + \| x \|_2$. Then the definition of the norm $\| x \|$ cannot be extended to include any sequences outside of the common linear space, since for any such sequence the l.u.b. expressions which define $\| x \|_1$ and $\| x \|_2$ are not finite. Therefore by the general completeness proof mentioned above, $\| x \|$ is a Banach norm. Since $\| x \|_1 \leqq \| x \|$ and $\| x \|_2 \leqq \| x \|$, the theorem now follows from a remark in the introduction, based on Theorem 5, p. 41 of Banach [2].

A norm $\| x \|$ defined on the linear space S_f may have one or both of the following properties:

(1) For every $x \epsilon S_f$,

$$\| x \| = \| (x_1, \cdots, x_n) \| = \| (| x_1 |, \cdots, | x_n |) \|.$$

(2) For every $x \epsilon S_f$, for each permutation $(n_1, n_2, \cdots, n_n)$ of $(1, 2, \cdots, n)$,

$$\| (x_1, x_2, \cdots, x_n) \| = \| (x_{n_1}, x_{n_2}, \cdots, x_{n_n}) \|.$$

The norm $\| x \|$ in a Minkowski space l_n of sequences $x = (x_1, \cdots, x_n)$ may have corresponding properties (fixed n in case of l_n). In case a norm has property (1), it will be called *symmetric*; if it has property (2), it will be called *permutable*. The Banach space S arising from a *symmetric* or *permutable* norm on S_f will correspondingly be called a *symmetric* or *permutable* sequence space.

In any sequence space S, the subspace of sequences in which all coördinates are zero except those in the $i_1, i_2, \cdots, i_n$ positions is an n-dimensional Minkowski space. We shall use the notation $l(i_1, i_2, \cdots, i_n)$ to denote such a subspace.

LEMMA 4.1. *For any sequence $\{x_i\}$ of a symmetric sequence space, or of a space S_f or l_n with a symmetric norm, if the numerical value of one coördinate is increased, the norm does not decrease. Thus there exists a projection of norm* 1 *on any Minkowski subspace $l(i_1, i_2, \cdots, i_n)$.*

Proof. Let $x = (x_1, x_2, \cdots, x_n)$ be any sequence such that $x_i \geqq 0$ for all i. Suppose $\Delta x_1 > 0$. If $x_1 > 0$, let

$$b = \Delta x_1(2x_1 + \Delta x_1)^{-1}, \qquad c = 2x_1(2x_1 + \Delta x_1)^{-1},$$

$$x' = (-x_1, x_2, \cdots, x_n), \qquad x + \Delta x = (x_1 + \Delta x_1, x_2, \cdots, x_n).$$

Then $b + c = 1$, $-bx_1 + c(x_1 + \Delta x_1) = x_1$, and $bx' + c(x + \Delta x) = x$. Thus

$$(b + c) \| x \| = \| bx' + c(x + \Delta x) \| \leqq b \| x' \| + c \| x + \Delta x \|;$$

or, since $\| x \| = \| x' \|$, $\| x \| \leqq \| x + \Delta x \|$. If $x_1 = 0$,

$$2 \| (0, x_2, \cdots, x_n) \| \leqq \| (\Delta x_1, x_2, \cdots, x_n) \| + \| (-\Delta x_1, x_2, \cdots, x_n) \|$$
$$= 2 \| (\Delta x_1, x_2, \cdots, x_n) \|.$$

This is sufficient to verify the lemma. The projection of norm 1 is the projection obtained by taking as the image of each sequence $x = \{x_i\}$ the point $(x_{i_1}, \cdots, x_{i_n})$ of the Minkowski space.

By Lemma 4.1, for any symmetric norm $\| x \|'$ on S_f, the second norm $\| x \|$ in the definition of S coincides with the given norm on S_f; and for any $x = \{x_i\}$ of S, $\lim_{n \to \infty} \| \{x_i\}_n \|$ exists, and $\| x \| = \lim_{n \to \infty} \| \{x_i\}_n \|$. Moreover, every sequence $x = \{x_i\}$ such that this limit exists belongs to S (by the definition of S).

A Banach space B is said to admit a base ([2], p. 110) when there exists a countable sequence $\{X_i\}$ of elements of the space such that every element $x \in B$ may be uniquely represented in the form

$$x = \sum_{i=1}^{\infty} x_i X_i \qquad \text{(i.e., if } s_n = \sum_{i=1}^{n} x_i X_i, \lim_{n \to \infty} \| x - s_n \| = 0).$$

By the triangle property of the norm, $| \| x \| - \| s_n \| | \leqq \| x - s_n \|$, and thus for every $x \in B$, $\| x \| = \lim_{n \to \infty} \| s_n \|$. Thus a space B with a base is essentially a sequence space which arises from a normed linear space S_f by the addition of certain infinite sequences $x = \{x_i\}$ such that $\lim_{n \to \infty} \| \{x_i\}_n \|$ exists, taking $\| x \|$ to be the value of this limit. If in B we define a second norm by $\| x \|_1 = \text{l.u.b.}_{1 \leqq n < \infty} \| \{x_i\}_n \|$, $\| x \|$ and $\| x \|_1$ are isomorphic on B.[15] Thus B is isomorphic to a closed linear subspace of our space S which would arise from the norm of B on S_f. If the norm in B is symmetric on S_f, then B may be identical with our space S, or equivalent to a (proper) closed linear subspace of S. The latter case occurs, for example, when the norm on S_f is $\| x \| = \max_i | x_i |$. The space S is then the space of bounded sequences $(m) = l_\infty$, while the space B is the space (c_0) of sequences convergent to 0.

For a Minkowski space l_n of sequences $x = (x_1, \cdots, x_n)$, or for a norm on S_f, obviously without loss of generality we may normalize so that

$$\| \{\delta_{ij}\} \| = 1 \qquad \text{for all } j.$$

In this case if the norm on S_f is symmetric, the sequences $X_i = \{\delta_{ij}\}$ are base elements of norm 1 for the smallest closed linear subspace B of S which contains the subspace S_f (i.e., B is the closure of S_f in S).

Suppose that we have a symmetric Minkowski space l_n of sequences $x =$

[15] It may easily be shown that B with the norm $\| x \|_1$ is complete, as is stated by Banach ([2], p. 111). The isomorphism then follows by Theorem 5, p. 41 of [2]. The space B with norm $\| x \|_1$ is closed in S, since obviously completeness and closure are equivalent for subspaces of a complete space.

$(x_1, \cdots, x_n)$, normalized as above. In addition to the given norm, let us introduce a Euclidean norm by $\| x \|_2 = (\sum x_i^2)^{\frac{1}{2}}$. The unit sphere for the Minkowski space is then the boundary of a convex region in the Euclidean space. Choose rectangular coördinate axes in the Euclidean space so that the coördinates of any sequence $(x_1, \cdots, x_n)$ are at the same time its rectangular coördinates. The unit sphere for the conjugate Minkowski space l_n is the set of all points $y = (y_1, \cdots, y_n)$ such that $y(x) = \sum_1^n y_i x_i$ is a functional of norm 1. By the definition of the conjugate space, we have Hölder's inequality

$$(y, x) \leqq \| y \|' \cdot \| x \|,$$

where $\| y \|'$ denotes the norm of the functional $y(x) = (y, x) = \sum_1^n y_i x_i$.

LEMMA 4.2. *Given any normalized symmetric Minkowski space of dimension* $n = 2^\nu$. *Let a Euclidean norm be introduced as above. Then there exists a subspace in which every involution has norm*

$$| U + V | \geqq \max \left[\frac{n^{\frac{1}{2}}}{\| X_1 \|}, \frac{n^{\frac{1}{2}}}{\| X_1 \|'} \right], \quad \textit{where } X_1 = (1, 1, \cdots, 1).$$

Proof. Let U be the orthogonal involution of §2, $U + V$ any other involution in the same subspace. Then, as in the proof of Theorem 2.1, since trace $(V) = 0$, we have for at least one k

$$1 \leqq 1 + v_{kk} = \sum_i u_{ik}(u_{ik} + v_{ik}) \leqq n^{-\frac{1}{2}} \| X_1 \|' \cdot \| (U + V) Y_k \|,$$

where $Y_k = \{\delta_{ik}\}$. Therefore

$$| U + V | \geqq \| (U + V) Y_k \| \geqq \frac{n^{\frac{1}{2}}}{\| X_1 \|'}.$$

Since trace $(VU) = 0$ by Theorem 1.2, for at least one j we have $\sum_i v_{ji}$ sign $u_{ji} \geqq 0$. If as in §3 we let $X_j = (\text{sign } u_{j1}, \cdots, \text{sign } u_{jn})$, then by Lemma 4.1, $\| (U + V) X_j \| \geqq | \sum_i v_{ji} \text{ sign } u_{ji} + n^{\frac{1}{2}} |$, and hence $\| (U + V) X_j \| \geqq n^{\frac{1}{2}}$. Therefore

$$| U + V | \geqq \frac{\| (U + V) X_j \|}{\| X_j \|} \geqq \frac{n^{\frac{1}{2}}}{\| X_j \|} = \frac{n^{\frac{1}{2}}}{\| X_1 \|}.$$

Suppose that we are given any normalized symmetric sequence space S. In addition to the given norm on the subspace S_f, let a Hilbert norm be introduced by $\| x \|_2 = (\sum x_i^2)^{\frac{1}{2}}$. Then in any Minkowski subspace $l(i_1, i_2, \cdots, i_n)$, the convex unit sphere has a *minimum* and *maximum* Euclidean radius. The corresponding directions in the Euclidean space will be called *minimal* and *maximal* directions.

THEOREM 4.2. *Given any symmetric sequence space S, which has the following properties:*

(1) *For any $x \epsilon S_f$, $\| x \| \leqq \| x \|_2$ (where $\| x \|_2$ is the Hilbert norm on S_f introduced above).*

(2) *For any n and any choice of i_k's, in the Minkowski space $l(i_1, i_2, \cdots, i_n)$, $X_1 = (1, 1, \cdots, 1)$ is the maximal direction.*

Then, if a projection exists on every closed linear subspace of S, S must be isomorphic to Hilbert space.

Similarly if $\| x \| \geqq \| x \|_2$ and if $(1, 1, \cdots, 1)$ is the minimal direction.

Proof. If $\| x \|$ and $\| x \|_2$ are isomorphic on S_f, the space S is isomorphic to Hilbert space (by the remark preceding Theorem 4.1). Suppose that $\| x \|$ and $\| x \|_2$ are not isomorphic. We divide the coördinates $(x_1, \cdots, x_n, \cdots)$ into blocks as follows: S_1 consists of x_1, x_2; S_2 consists of x_3 through x_{2+n_2}, where $n_2 = 2^{\nu_2}$, ν_2 being chosen sufficiently large so that in the space $l(3, \cdots, 2 + n_2)$, $(n_2)^{\frac{1}{2}}/\| X_{1,n_2} \| \geqq 2$, where $X_{1,n_2} = (1, 1, \cdots, 1)$; $\cdots$; S_k consists of $x_{3+n_2+\cdots+n_{k-1}}$ through $x_{2+n_2+\cdots+n_k}$, where $n_k = 2^{\nu_k}$, ν_k being chosen sufficiently large so that in the space $l(3 + n_2 + \cdots + n_{k-1}, \cdots, 2 + n_2 + \cdots + n_k)$, $(n_k)^{\frac{1}{2}}/\| X_{1,n_k} \| \geqq k$, where $X_{1,n_k} = (1, 1, \cdots, 1)$; $\cdots$. This is possible since if $\| x \|$ and $\| x \|_2$ are not isomorphic on S_f, they are not isomorphic on any subspace of S_f obtained by omitting a finite number of coördinates.

We now define a closed linear subspace l of S as follows. The elements of l are those sequences of S whose coördinates in the block S_k, for every k, are coördinates of the extreme subspace determined by the matrix β_{n_k}. There exists no projection on this subspace.[16] For by Lemma 4.1, there exists a projection of norm 1 of S on the Minkowski subspace of the coördinates in S_k, so that for each k we may apply Lemma 2.1. By Lemma 4.2 and the construction of the S_k's, this yields $C(l) \geqq k$ for each k, or $C(l) = \infty$.

A similar construction and argument may be given for the case $\| x \| \geqq \| x \|_2$ and $(1, 1, \cdots, 1)$ the *minimal* direction.

The hypotheses of Theorem 4.2 evidently may be weakened. For any symmetric sequence space, let $r_{1,n}$ denote the Euclidean radius of the convex unit sphere of $l(1, 2, \cdots, n)$ in the $(1, 1, \cdots, 1)$ direction: and let $r'_{1,n}$ denote the Euclidean radius in the $(1, 1, \cdots, 1)$ direction of the conjugate Minkowski space to $l(1, 2, \cdots, n)$, as in Lemma 4.2. Then if $y(x) = a \sum_1^n x_i$ is a functional of norm 1 on $l(1, 2, \cdots, n)$, $1/r'_{1,n}$ = perpendicular Euclidean distance from the origin to the hyperplane $y(x) = 1$. We may now state the following theorem.

THEOREM 4.3. *In any symmetric sequence space, if there exists a projection on every subspace, it is necessary that $n^{\frac{1}{2}}/\| X_{1,n} \| = r_{1,n}$ be bounded as $n \rightarrow \infty$, where*

[16] If B is the closure of S_f in S, or any space intermediate between $\bar{S}_f$ and S, it follows by this same argument that there is a closed linear subspace $l \subset B$ such that there exists no projection of B on l. Cf. footnote 8.

$X_{1,n} = (1, 1, \cdots, 1)$ *in* $l(1, 2, \cdots, n)$. *It is also necessary that* $n^{\frac{1}{2}}/\| X_{1,n} \|' = r'_{1,n}$ *be bounded as* $n \to \infty$.

Proof. Suppose that $n^{\frac{1}{2}}/\| X_{1,n} \|$ is unbounded. For any fixed $h > 1$, let $X_{1,h+n} = X_{1,h} + X_{h,n}$, where $X_{h,n} = (1, 1, \cdots, 1)$ in $l(h + 1, \cdots, h + n)$. Then by Lemma 4.1

$$\frac{(n+h)^{\frac{1}{2}}}{\| X_{1,n+h} \|} \leqq \frac{(n+h)^{\frac{1}{2}}}{\| X_{h,n} \|} = \frac{(n+h)^{\frac{1}{2}}}{n^{\frac{1}{2}}} \cdot \frac{n^{\frac{1}{2}}}{\| X_{h,n} \|},$$

so that $n^{\frac{1}{2}}/\| X_{h,n} \|$ is unbounded. Thus for any k, there is an n such that $n^{\frac{1}{2}}/\| X_{h,n} \| > 2^{\frac{1}{2}} \cdot k$. Let ν_k be such that $n_k = 2^{\nu_k} \leqq n \leqq 2n_k$. Then

$$\frac{(n_k)^{\frac{1}{2}}}{\| X_{h,n_k} \|} = \left(\frac{n_k}{n}\right)^{\frac{1}{2}} \frac{n^{\frac{1}{2}}}{\| X_{h,n_k} \|} \geqq \left(\frac{n_k}{n}\right)^{\frac{1}{2}} \frac{n^{\frac{1}{2}}}{\| X_{h,n} \|} > \left(\frac{n_k}{n}\right)^{\frac{1}{2}} \cdot 2^{\frac{1}{2}} \cdot k \geqq k$$

by Lemma 4.1. In this argument, first choose $h = 1$ and $k = 1$. This yields an n_1 such that $n_1^{\frac{1}{2}}/\| X_{1,n_1} \| \geqq 1$. Then, choosing $h = n_1$ and $k = 2$, we obtain an n_2 such that $n_2^{\frac{1}{2}}/\| X_{n_1,n_2} \| \geqq 2$. By induction we see that there is a sequence $\{n_k\}$ such that, for each k, $n_k^{\frac{1}{2}}/\| X_{h_k,n_k} \| \geqq k$, where $h_k = n_1 + \cdots + n_{k-1}$. Therefore, as in the proof of Theorem 4.2, we may construct a subspace on which there exists no projection.

The second statement of Theorem 4.3 follows by Lemma 1.3.

For permutable symmetric sequence spaces, $1/r'_{1,n} = r_{1,n}$.[17] Thus for these spaces, Theorem 4.3 states that for the existence of projections it is necessary that the Euclidean radius of the convex unit sphere in the $(1, 1, \cdots, 1)$ direction be bounded both from 0 and from ∞ as $n \to \infty$.

5. **Spaces defined by a sequence of two-dimensional norms.** We consider in this section a class of symmetric sequence spaces which arise from norms on S_f defined as follows. Let $\| x_1, x_2 \|_2$, $\| x_1, x_2 \|_3, \cdots$ be any sequence of symmetric two-dimensional norms. Then a symmetric norm on S_f is defined by $\| (x_1, x_2) \| = \| (x_1, x_2) \|_2$, $\| (x_1, x_2, x_3) \| = \| \| (x_1, x_2) \|_2, x_3 \|_3$, $\| (x_1, x_2, x_3, x_4) \| = \| \| \| (x_1, x_2) \|_2, x_3 \|_3, x_4 \|_4$, etc. A norm for a Minkowski space l_n may be similarly defined, given any $n - 1$ two-dimensional norms.

Lemma 5.1. *Given any Minkowski space* l_n *defined by a sequence* $\{\| x_1, x_2 \|_i\}$ *of two-dimensional norms* $(i = 2, 3, \cdots, n)$. *The conjugate Minkowski space* l_n *is equivalent to the space defined by the sequence* $\{\| x_1, x_2 \|'_i\}$ $(i = 2, 3, \cdots, n)$, *where, for each* i, $\| x_1, x_2 \|'_i$ *is the conjugate norm to* $\| x_1, x_2 \|_i$.

Proof. The proof is by induction. By Hölder's inequality, as in Lemma 4.2 if $\sum_1^n y_i x_i$ is any functional on l_n,

$$(y_1 x_1 + y_2 x_2) \leqq \| y_1, y_2 \|' \cdot \| x_1, x_2 \|;$$

[17] For in these spaces, if $y(x) = a \sum_1^n x_i$ and $b \cdot X_{1,n}$ are of norm 1, then obviously the hyperplane $y(x) = 1$ has the point $b \cdot X_{1,n}$ for a point of contact with the unit sphere.

and if

$$\sum_1^{n-1} y_i x_i$$

$$\leqq \| \cdots \| \| y_1, y_2 \|_2', y_3 \|_3', \cdots, y_{n-1} \|_{n-1}' \cdot \| \cdots \| \| x_1, x_2 \|_2, x_3 \|_3, \cdots, x_{n-1} \|_{n-1},$$

then

$$\sum_1^n y_i x_i \leqq \| \cdots \| \| y_1, y_2 \|_2', y_3 \|_3', \cdots, y_{n-1} \|_{n-1}'$$
$$\cdot \| \cdots \| \| x_1, x_2 \|_2, x_3 \|_3, \cdots, x_{n-1} \|_{n-1} + y_n x_n$$
$$\leqq \| \cdots \| \| y_1, y_2 \|_2', y_3 \|_3', \cdots, y_n \|_n' \cdot \| \cdots \| \| x_1, x_2 \|_2, x_3 \|_3, \cdots, x_n \|_n.$$

The lemma follows since, for any $\{y_i\}$, obviously there exists $\{x_i\}$ such that equality is attained in Hölder's inequality.

An important particular case is that of the $p_2p_3p_4 \cdots$ spaces.[18] Such a space is the Banach space S obtained from the norm on S_f defined as above, with $\| x_1, x_2 \|_n = (| x_1 |^{p_n} + | x_2 |^{p_n})^{1/p_n} = \| x_1, x_2 \|_{p_n}$ for each n, where $\{p_n\}$ is any sequence such that $\infty \geqq p_n \geqq 1$ for all n. In any $p_2p_3p_4 \cdots$ space we define a Euclidean norm $\| x \|_2$ on the linear subspace S_f as usual by

$$\| x \|_2 = \left(\sum_1^n x_i^2\right)^{\frac{1}{2}}.$$

LEMMA 5.11. *If in a $p_2p_3p_4 \cdots$ space the p_i's all satisfy $p_i \leqq 2$, then*

$$\| x \|_2 \leqq \| x \| = \| x \|_{p_2p_3p_4 \cdots}.$$

If $p_i \geqq 2$ for all i, then

$$\| x \|_2 \geqq \| x \| = \| x \|_{p_2p_3p_4 \cdots}.$$

The proof of the lemma is by induction. We use the inequality

$$\left(\sum a_i^{p_2}\right)^{1/p_2} \leqq \left(\sum a_i^{p_1}\right)^{1/p_1}, \qquad p_1 \leqq p_2, \; a_i \geqq 0.$$ [19]

Suppose $p_i \leqq 2$. Then, as a special case of the inequality, $\| x_1, x_2 \|_2 \leqq \| x_1, x_2 \|_{p_i}$ for each i, and

$$\| \| x_1, x_2 \|_2, x_3 \|_2 \leqq \| \| x_1, x_2 \|_2, x_3 \|_{p_3} \leqq \| \| x_1, x_2 \|_{p_2}, x_3 \|_{p_3}$$

by Lemma 4.1. If

$$\| \cdots \| x_1, x_2 \|_2 \cdots, x_{n-1} \|_2 \leqq \| \cdots \| x_1, x_2 \|_{p_2} \cdots, x_{n-1} \|_{p_{n-1}},$$

then, by the special case of the inequality and Lemma 4.1,

$$\| \| \cdots \| x_1, x_2 \|_2 \cdots, x_{n-1} \|_2, x_n \|_2 \leqq \| \| \cdots \| x_1, x_2 \|_2 \cdots, x_{n-1} \|_2, x_n \|_{p_n}$$
$$\leqq \| \| \cdots \| x_1, x_2 \|_{p_2} \cdots, x_{n-1} \|_{p_{n-1}}, x_n \|_{p_n}.$$

[18] The norm for the $p_2p_3p_4 \cdots$ spaces was suggested to me by Dr. J. W. Tukey.

[19] A reference for this inequality is given in footnote 10.

This proves the first statement of the lemma. The proof for the case of the p_i's $\geqq 2$ is similar, with the inequality signs reversed.

We consider first the $p_2p_3p_4 \cdots$ spaces where the $p_i \leqq 2$. For such a $p_2p_3p_4 \cdots$ space, there will be sequences $\{C_n\}$ of constants such that in the Minkowski subspaces $l(1, 2, \cdots, n)$, for each n,

$$\| x \|_2 \leqq \| x \|_{p_2 \cdots p_n} \leqq C_n \| x \|_2$$

for all $x \in l(1, 2, \cdots, n)$.

LEMMA 5.12. *If there exists a set of C_n's which are bounded as $n \to \infty$, the $p_2p_3p_4 \cdots$ space is isomorphic to Hilbert space.*

This lemma follows from the remark preceding Theorem 4.1.

LEMMA 5.2. *In any $p_2p_3p_4 \cdots$ space, $p_i \leqq 2$, the smallest possible C_n's satisfy the difference system*

$$C_n^{s_n} - C_{n-1}^{s_n} = 1, \qquad C_1 = 1, \qquad\qquad \textit{where } \frac{1}{s_n} = \frac{1}{p_n} - \frac{1}{2}.$$

Proof. By Hölder's inequality $(a_1b_1 + a_2b_2) \leqq (a_1^p + a_2^p)^{1/p} \cdot (b_1^q + b_2^q)^{1/q}$ or $(a, b) \leqq \| a \|_p \cdot \| b \|_q$ $(p^{-1} + q^{-1} = 1)$, if $p = 2/p_2$, $(x_1^{p_2} \cdot 1 + x_2^{p_2} \cdot 1)^{1/p_2} \leqq (x_1^2 + x_2^2)^{\frac{1}{2}} \cdot 2^{1/p_2 - \frac{1}{2}} = (x_1^2 + x_2^2)^{\frac{1}{2}} \cdot C_2$, $C_2 = 2^{1/s_2} = \| 1, 1 \|_{s_2}$, and if $2/p_{k+1} = p$,

$$\{\{\cdots\{x_1^{p_2} + \cdots + x_k^{p_k}\}^{p_{k+1}/p_k} + x_{k+1}^{p_{k+1}}\}^{1/p_{k+1}}$$

$$\leqq \{C_k^{p_{k+1}} \| x \|_{2k}^{p_{k+1}} + x_{k+1}^{p_{k+1}}\}^{1/p_{k+1}}$$

$$\leqq (\| x \|_{2k}^2 + x_{k+1}^2)^{\frac{1}{2}p_{k+1} \cdot 1/p_{k+1}} \cdot (C_k^{p_{k+1} \cdot 2/(2-p_{k+1})} + 1)^{(1-\frac{1}{2}p_{k+1})/p_{k+1}}$$

$$= (C_k^{s_{k+1}} + 1)^{1/s_{k+1}} \cdot \| x \|_{2(k+1)} = C_{k+1} \cdot \| x \|_{2(k+1)},$$

where $\| x \|_{2k} = \| \{x_i\}_k \|_2$. Thus

$$C_3 = (C_2^{s_3} + 1)^{1/s_3} = [\| 1, 1 \|_{s_2}^{s_3} + 1]^{1/s_3} = \| 1, 1, 1 \|_{s_2 s_3},$$

$$C_{k+1} = (\| 1, 1, \cdots, 1 \|_{s_2 \cdots s_k}^{s_{k+1}} + 1)^{1/s_{k+1}} = \| 1, 1, \cdots, 1 \|_{s_2 \cdots s_{k+1}},$$

and for all n, C_n is given by $C_n^{s_n} - C_{n-1}^{s_n} = 1$, if $C_1 = 1$. For each k, C_{k+1} as above is the least possible constant, since at each stage there is a value of x_{k+1} for which equality is attained in the Hölder's inequality.

LEMMA 5.3. *A n.s.c. that the C_n's in the solution $(1, C_2, C_3, \cdots)$ of the difference system*

$$C_n^{s_n} - C_{n-1}^{s_n} = 1$$

be bounded is that there exist a constant K such that the series $\sum_2^\infty K^{-s_n}$ is convergent. A sufficient condition is that there exist $\epsilon > 0$ such that, for all n, $s_n/\log n > \epsilon$. If the p_n's (and therefore the s_n's) are monotone, $1 \leqq p_2 \leqq p_3 \leqq \cdots \leqq 2$, the latter condition is also necessary.

Proof. Let $C_n = \exp(c_n)$, $c_n = \log C_n$, and let $s_n = w_n \log n$. Then

$$\exp(c_n \cdot w_n \log n) - \exp(c_{n-1} \cdot w_n \log n) = n^{c_n \cdot w_n} - n^{c_{n-1} \cdot w_n} = 1.$$

By the mean value theorem (since $dn^x/dx = n^x \log n$),

$$n^{\bar{x}} \cdot \log n \cdot (c_n w_n - c_{n-1} w_n) = 1$$

or

$$c_n - c_{n-1} = \frac{1}{n^{\bar{x}} \cdot w_n \cdot \log n} = \frac{1}{n^{\bar{x}} \cdot s_n}, \quad \text{where } c_{n-1} w_n \leqq \bar{x} \leqq c_n w_n .$$

Summing the members of the last equality, we have

$$c_k = \sum_2^k (c_n - c_{n-1}) = \sum_2^k \frac{1}{n^{\bar{x}} \cdot s_n} .$$

Also

$$n^{c_{n-1} w_n} = C_{n-1}^{s_n} \leqq n^{\bar{x}} \leqq n^{c_n w_n} = C_n^{s_n},$$

$$\sum_2^k \frac{1}{C_n^{s_n} \cdot s_n} \leqq \sum_2^k \frac{1}{n^{\bar{x}} \cdot s_n} = c_k \leqq \sum_2^k \frac{1}{C_{n-1}^{s_n} \cdot s_n} .$$

If the C_k's are bounded, so are the c_k's, and this inequality implies that $\sum_2^\infty (K^{s_n} \cdot s_n)^{-1}$ is convergent for some K. Conversely if $\sum (K^{s_n} \cdot s_n)^{-1}$ is convergent for some K, the C_n's are bounded. For suppose they are unbounded. Then by the second half of the last inequality, $\sum_2^\infty (C_{n-1}^{s_n} \cdot s_n)^{-1}$ is divergent. But also by hypothesis, for a sufficiently large N,

$$\sum_N^\infty \frac{1}{K^{s_n} \cdot s_n} > \sum_N^\infty \frac{1}{C_{n-1}^{s_n} \cdot s_n}$$

(the C_n's are monotone), and we have a contradiction.

Since $s_n \geqq 2$ and $2^{s_n} > s_n$ for all n, we have

$$\sum \frac{1}{K^{s_n} \cdot s_n} < \sum \frac{1}{K^{s_n}} \qquad \text{and} \qquad \sum \frac{1}{(2K)^{s_n}} < \sum \frac{1}{K^{s_n} \cdot s_n} .$$

Therefore the condition that there exist a K such that $\sum (K^{s_n} \cdot s_n)^{-1}$ is convergent is equivalent to the condition that there exist a K such that $\sum K^{-s_n}$ is convergent.

To verify the second statement of Lemma 5.3, suppose $s_n / \log n > \epsilon > 0$ for all n, and that the c_n's are unbounded. Then since $w_n > \epsilon$ for all n ($s_n = w_n \log n$), and $c_{n-1} w_n \leqq \bar{x}$, for sufficiently large n we have $\bar{x} > 2$. Therefore $c_n - c_{n-1} = (n^{\bar{x}} \cdot s_n)^{-1} < n^{-2}$, and $c_k = \sum_2^k (c_n - c_{n-1}) < \sum_2^k n^{-2}$. This is a contradiction, since $\sum_2^\infty n^{-2}$ is convergent.

Finally, in case the p_n's are monotone, we have $2 \leqq s_2 \leqq s_3 \leqq \cdots \leqq \infty$. Let us compare the solutions of the system $C_n^{s_n} - C_{n-1}^{s_n} = 1$ with those of the system $D_n^{s_n} - D_{n-1}^{s_{n-1}} = 1$ ($D_1 = 1$, $s_1 = s_2$). Obviously $D_n \leqq C_n$ for all n, and $D_n^{s_n} = n$. Therefore if the C_n's are bounded, $n^{1/s_n} = D_n \leqq C_n < K$, or $s_n^{-1} \log n < \log K$, $s_n/\log n > 1/\log K$ for all n.

Examples may be easily constructed to show that the condition $s_n/\log n > \epsilon > 0$ is not necessary in the general non-monotone case.[20]

In the following theorem we do not assume that $p_i \leqq 2$.

THEOREM 5.1. *Given any $p_2p_3p_4 \cdots$ space. A necessary and sufficient condition that a sequence $x = (x_1, x_2, \cdots)$ belong to the space is that $\sum_2^\infty x_n^{p_n}/(K^{p_n}\cdot p_n)$ be convergent for some K. Thus the given $p_2p_3p_4 \cdots$ space and any other $p_2p_3p_4 \cdots$ space obtained by rearranging the p_n's in any manner contain the same sequences correspondingly rearranged.*

Proof. The proof is similar to the proof of Lemma 5.3. With any sequence $x = (x_1, x_2, \cdots)$ we associate the difference system $C_n^{p_n} - C_{n-1}^{p_n} = x_n^{p_n}$, $C_1 = x_1$. Obviously the C_n's are monotone; and if x belongs to the $p_2p_3p_4 \cdots$ space, $\lim_{n\to\infty} C_n = \| x \|$. Let $C_n = \exp(c_n)$; in the same way as in the proof of Lemma 5.3 we obtain

$$\frac{x_n^{p_n}}{C_n^{p_n}\cdot p_n} < c_n - c_{n-1} = \frac{x_n^{p_n}}{n^{\bar{x}}\cdot p_n} < \frac{x_n^{p_n}}{C_{n-1}^{p_n}\cdot p_n}, \quad c_{n-1}\frac{p_n}{\log n} \leqq \bar{x} \leqq c_n \frac{p_n}{\log n}.$$

The first statement of Theorem 5.1 follows from this. The given $p_2p_3p_4 \cdots$ space and any rearranged $p_2p_3p_4 \cdots$ space of course contain the same sequences correspondingly rearranged since any series of positive terms may be rearranged in any manner without affecting the convergence or sum of the series.

COROLLARY TO THEOREM 5.1. *Given a $p_2p_3p_4 \cdots$ space. Then any rearranged $p_2p_3p_4 \cdots$ space is isomorphic to the given space. The isomorphism is given by the identity transformation in the common linear space of sequences.*

This corollary is a consequence of Theorem 4.1.

THEOREM 5.12. *Given any $p_2p_3p_4 \cdots$ space. The space contains those and only those sequences $x = \{x_i\}$ for which there exists a $k > 0$ such that $\sum_{n=1}^\infty (kx_n)^{p_n}$*

[20] One such example is as follows. Choose a subsequence $\{n_k\}$ of $\{n\}$ such that

$$\sum_{k=1}^\infty \frac{1}{\log n_k} = A < \infty.$$

Let $\{n\} - \{n_k\}$ be $\{m_j\}$. Define $s_{n_k} = \log \log n_k$, $s_{m_j} = 2 \log m_j$. Then if $K = e$,

$$\sum_n \frac{1}{K^{s_n}} = \sum_k \frac{1}{\log n_k} + \sum_j \frac{1}{m_j^2} < A + \sum_m \frac{1}{m^2} < \infty,$$

so the C_n's are bounded. But since $s_{n_k} = \log \log n_k = w_k \log n_k$, $w_k = s_{n_k}/\log n_k \to 0$.

is convergent. If the p_n's are bounded, the space contains those and only those sequences $x = \{x_i\}$ such that $\sum_1^\infty x_n^{p_n}$ is convergent. The latter characterization is not sufficient in any case when the p_n's are unbounded.

Proof. Suppose a sequence $\{x_i\}$ is such that $\sum_n (kx_n)^{p_n}$ is convergent. Then if $K = k^{-1}$, $\sum_n x_n^{p_n}/(K^{p_n}\cdot p_n)$ is convergent. If $\sum_n x_n^{p_n}/(K^{p_n}\cdot p_n)$ is convergent, $\sum_n x_n^{p_n}(2K)^{-p_n}$ is convergent since $2^{p_i} > p_i$. This proves the first statement, by Theorem 5.1.

If the p_n's are bounded, obviously $\sum_n (kx_n)^{p_n} < \infty$ implies $\sum_n x_n^{p_n} < \infty$. Suppose the p_n's are unbounded. Choose a subsequence $\{p_{n(k)}\}$ such that $p_{n(k)} \geqq \log k/\log 2$. Define a sequence $\{a_n\}$ by

$$a_{n(k)} = \frac{1}{k^{1/p_{n(k)}}\cdot 2}, \qquad a_n = 0 \text{ otherwise.}$$

Then $2^{p_{n(k)}} \geqq 2^{\log k/\log 2} = \exp(\log k) = k$, so that

$$a_{n(k)}^{p_{n(k)}} \leqq \frac{1}{k^2}, \qquad \text{and} \qquad \sum_n a_n^{p_n} = \sum_k a_{n(k)}^{p_{n(k)}} \leqq \sum \frac{1}{k^2} < \infty.$$

But $\sum_n (2a_n)^{p_n} = \sum_k k^{-1} = \infty$. Thus the sequence $x = \{2a_n\}$ satisfies the condition $\sum_n (kx_n)^{p_n} < \infty$ (with $k = \frac{1}{2}$), but not the condition $\sum_n x_n^{p_n} < \infty$. This verifies the last sentence of the theorem.

The sequences $\{x_i\}$ such that $\sum_{n=1}^\infty x_n^{p_n}$ is convergent have been studied by W. Orlicz [7]. Orlicz does not recognize, however, the possibility of introducing a norm so that the class of sequences becomes a Banach space.

THEOREM 5.121. *Given a $p_2p_3p_4 \cdots$ space, such that $1 \leqq p < p_n$ for all n. A n.s.c. that the $p_2p_3p_4 \cdots$ space be isomorphic to l_p is that there exist a constant k, $0 < k < 1$, such that $\sum_{n=1}^\infty k^{r_n}$ is convergent, where $r_n = p_n/(p_n - p)$.*

This theorem follows by Theorem 4.1 and a theorem of Orlicz ([7], Satz 3, p. 205).

THEOREM 5.122. *A n.s.c. that a $p_2p_3p_4 \cdots$ space be isomorphic to $(m) = l_\infty$ is that there exist a $k > 0$ such that $\sum_n k^{p_n}$ is convergent. If the p_n's are monotone, this condition is that there exist $\epsilon > 0$ such that for all n, $p_n/\log n > \epsilon$* (as in Lemma 5.3).

This theorem follows by Theorems 5.12 and 4.1.

LEMMA 5.4. *Given any $p_2p_3p_4 \cdots$ space (some p_i's > 2 and some $\leqq 2$). Let those p_i's which > 2 be relabeled $q_2, q_3, \cdots$; and let those which $\leqq 2$ be labeled $p_2, p_3, \cdots$. The given space is isomorphic to Hilbert space if (and only if) both*

the $p_2p_3p_4 \cdots$ subspace and the $q_2q_3q_4 \cdots$ subspace are Hilbert spaces (or one Hilbert space and the other finite dimensional).

Proof. If in the given space the first exponent $p_2 \leqq 2$, then by the $p_2p_3p_4 \cdots$ subspace we mean the subspace of sequences $x = (x_1, x_2, \cdots)$ in which all coördinates are 0 except those which correspond to $p_2, p_3, \cdots$, and the one coördinate which precedes p_2; by the $q_2q_3q_4 \cdots$ subspace we mean the subspace of sequences in which all coördinates are 0 except those which correspond to $q_2, q_3, \cdots$; and similarly if in the given space the first exponent > 2. Let x_p denote a sequence of the $p_2p_3p_4 \cdots$ subspace, x_q a sequence of the $q_2q_3q_4 \cdots$ subspace. Then any sequence x of the given space may be expressed as $x = x_p + x_q$. By hypothesis we have $\| x_p \|_2 \leqq \| x_p \| \leqq K \cdot \| x_p \|_2$ and $\| x_q \|_2 \geqq \| x_q \| \geqq k \cdot \| x_q \|_2$. Obviously $\| x_p \| + \| x_q \| \geqq \| x_p + x_q \| \geqq \max (\| x_p \|, \| x_q \|)$. Therefore for any finite sequence x we have

$$2^{\frac{1}{2}}K\| x_p + x_q \|_2 = 2^{\frac{1}{2}}K(\| x_p \|_2^2 + \| x_q \|_2^2)^{\frac{1}{2}} \geqq K(\| x_p \|_2 + \| x_q \|_2)$$
$$\geqq K\| x_p \|_2 + \| x_q \|_2 \geqq \| x_p \| + \| x_q \| \geqq \| x_p + x_q \| \geqq \tfrac{1}{2}(\| x_p \| + \| x_q \|)$$
$$\geqq \tfrac{1}{2}(\| x_p \|_2 + k \cdot \| x_q \|_2) \geqq \tfrac{1}{2}k(\| x_p \|_2^2 + \| x_q \|_2^2)^{\frac{1}{2}} = \tfrac{1}{2}k\| x_p + x_q \|_2$$

or

$$2^{\frac{1}{2}}K\| x_p + x_q \|_2 \geqq \| x_p + x_q \| \geqq \tfrac{1}{2}k\| x_p + x_q \|_2 ;$$

and by the remark preceding Theorem 4.1 the same inequalities are true for any infinite sequence x of the given space. (Conversely, if the given space is isomorphic to Hilbert space, the $p_2p_3p_4 \cdots$ and $q_2q_3q_4 \cdots$ subspaces are isomorphic to Hilbert or Euclidean spaces, since they are closed linear subspaces of the given space, and since, as is known, any closed linear subspace of a Hilbert space is either a Hilbert or a Euclidean space.[21]) This verifies Lemma 5.4.

Lemmas 5.1, 5.3 and 5.4 imply that any $p_2p_3p_4 \cdots$ space is isomorphic to Hilbert space if there exists $k > 0$ such that $\sum_1^\infty k^{s_i}$ is convergent, where $s_i^{-1} = | p_i^{-1} - \frac{1}{2} |$. ($k^{s_i} = 0$ if $p_i = 2$.)

[21] The question arises here as to whether a $p_2p_3p_4 \cdots$ space may be isomorphic to Hilbert space, the 1-1 correspondence for the isomorphism being some other than the identity correspondence of sequences of the $p_2p_3p_4 \cdots$ space and of the space $l_{2,\infty}$, while at the same time the identity correspondence is not an isomorphism (or cannot be defined because, no matter how the coördinates in the $p_2p_3p_4 \cdots$ space are rearranged, the $p_2p_3p_4 \cdots$ space and $l_{2,\infty}$ contain different sequences). Later results of this section show that the answer is negative (since if such a situation were possible we could find a $p_2p_3p_4 \cdots$ space having a projection for every closed linear subspace, but such that $\| x \|_{p_2p_3 \cdots}$ and $\| x \|_2$ would not be isomorphic on the subspace S_f). In fact it may be shown that if $p_i \leqq 2$ or $p_i \geqq 2$ for all i, and if the $p_2p_3p_4 \cdots$ space is isomorphic to Hilbert space, then the identity correspondence defines an isomorphism for which the affine ratio is a minimum. Strictly, however, in the present discussion "isomorphic" in Lemma 5.4 means isomorphic in the restricted sense that the identity correspondence is the isomorphism.

THEOREM 5.2. *Given any $p_2p_3p_4 \cdots$ space. If there exists a projection on every closed linear subspace, the space must be isomorphic to Hilbert space.*

Proof. We consider the $p_2p_3p_4 \cdots$ and the $q_2q_3q_4 \cdots$ subspaces, as in the proof of Lemma 5.4. If projections exist on every subspace of the given space, the p_i's and the q_i's must *converge* to 2. For if there were an infinite number of p_i's $\leqq p < 2$, or an infinite number of q_i's $\geqq q > 2$, in the subspace determined by these p_i's or q_i's we should have

$$\| x \|_{p_i} \geqq \| x \|_p \qquad \text{or} \qquad \| x \|_{q_i} \leqq \| x \|_q ,$$

and therefore

$$\frac{n^{\frac{1}{2}}}{\| X_1 \|_{q_i}} \geqq \frac{n^{\frac{1}{2}}}{\| X_1 \|_q} \qquad \left(\text{for } p_i\text{'s}, \frac{1}{q_i} = 1 - \frac{1}{p_i}\right),$$

so that by Theorem 4.3 we could construct a subspace on which there could exist no projection. If $p_n = 2$, $C_{n+1} = C_n$; so the $p_2p_3p_4 \cdots$ subspace is Hilbert space if and only if the subspace determined by the p_i's which < 2 is Hilbert space. Since $\lim_{n\to\infty} p_n = 2$, it is possible to rearrange the p_i's which < 2 in monotone order, $1 \leqq p_{n_1} \leqq p_{n_2} \leqq \cdots < 2$. By Lemmas 5.1, 1.3, and 5.4, and the corollary to Theorem 5.1, it is now sufficient to consider only the case of a $p_2p_3p_4 \cdots$ space, where $1 \leqq p_2 \leqq p_3 \leqq \cdots < 2$. In such a $p_2p_3p_4 \cdots$ space, suppose a projection exists on every subspace. Then by Theorem 4.3, it is necessary that

$$\frac{n^{\frac{1}{2}}}{\| X_{1,n} \|'} \geqq \frac{n^{\frac{1}{2}}}{\| X_{1,n} \|_{p_n'}} = n^{1/p_n - \frac{1}{2}}$$

be bounded as $n \to \infty$. But if $n^{1/p_n - \frac{1}{2}} \leqq K$, $(p_n^{-1} - \frac{1}{2}) \log n < \log K$, or

$$\frac{\log n}{s_n} < \log K, \qquad \frac{s_n}{\log n} > \log K \qquad \text{for all } n.$$

By Lemmas 5.12, 5.2, and 5.3, this requires the space to be Hilbert space.

By Theorem 5.12, the underlying linear space for any $p_2p_3p_4 \cdots$ space consists of those and only those sequences $x = \{x_i\}$ which are such that there exists $k > 0$ such that $\sum_2^\infty (kx_i)^{p_i} < \infty$. An alternative way of defining a norm for this linear space is as follows: The unit sphere consists of all sequences such that $\sum_1^\infty x_i^{p_i} = 1$. (Let $p_1 = p_2$. As usual $x_i^{p_i}$ means $| x_i |^{p_i}$.) By use of Hölder's inequality $(x + y)^p \leqq 2^{p-1}(x^p + y^p)$, it may be easily verified that this unit sphere is convex. The norm is defined for all sequences by the homogeneity property of the norm: if $x = \{x_i\}$ is any sequence of the space, there exists k such that $\sum_1^\infty (kx_i)^{p_i} = 1$, and $\| x \| = k^{-1}$. Consider this norm on the subspace S_f. If $x = \{x_i\}$ is any sequence such that $\lim_{n\to\infty} \| \{x_i\}_n \|$ is finite, then for each

n, $\| \{x_i\}_n \| = k_n^{-1}$, where $\| \{k_n x_i\}_n \| = 1 = \sum_{i=1}^{n} (k_n x_i)^{p_i}$, and l.u.b. k_n^{-1} is finite (lim = l.u.b. since the norm is symmetric). Therefore there exists a k, $0 < k < k_n$, and $\sum_{i=1}^{n} (kx_i)^{p_i} < 1$ for every n. Thus $\sum_{1}^{\infty} (kx_i)^{p_i} < \infty$, and $x = \{x_i\}$ belongs to the underlying linear space. Therefore, by the general completeness proof (§4), the new norm is a Banach norm, and is isomorphic to the $p_2p_3p_4 \cdots$ norm. In virtue of the isomorphism, we obtain

THEOREM 5.3. *In any $p_1p_2p_3 \cdots$ space of sequences, with norm*

$$\| x \| = \frac{1}{k}, \qquad \text{where } x = \{x_i\}, \sum_{1}^{\infty} (kx_i)^{p_i} = 1,$$

if a projection exists on every closed linear subspace, the space must be isomorphic to Hilbert space.

We return now to the consideration of spaces defined by a general sequence of two-dimensional norms $\| x, y \|^{(2)}, \| x, y \|^{(3)}, \cdots$. Suppose we are given such a space in which it is possible to construct, by the method of Theorem 4.3, a subspace on which there exists no projection. Then either $n^{\frac{1}{2}}/\| X_{1,n} \|$ or $n^{\frac{1}{2}}/\| X_{1,n} \|'$ is unbounded as $n \to \infty$. By Lemmas 1.3 and 5.1 we may assume without loss of generality that it is $n^{\frac{1}{2}}/\| X_{1,n} \|$ which is unbounded.

THEOREM 5.4. (*Comparison Theorem.*) *Suppose we are given a space defined by $\| x, y \|^{(2)}, \| x, y \|^{(3)}, \cdots$, the two-dimensional norms being normalized so that $\| 1, 0 \|^{(i)} = \| 0, 1 \|^{(i)} = 1$ for all i. Suppose further that we are given a second space defined by $\| x, y \|^{((2))}, \| x, y \|^{((3))}, \cdots$, where $\| 1, 0 \|^{((i))} = \| 0, 1 \|^{((i))} = 1$ for all i, which is such that there exists a constant c, such that for all i*

$$\| x, cy \|^{((i))} = \| x, y \|^{(i)\prime} \geqq \| x, y \|^{(i)}.$$

Then if by the method of Theorem 4.3 it is possible to construct a subspace in the second space on which there exists no projection (i.e., if $n^{\frac{1}{2}}/\| X_{1,n} \|^{((2,3,\cdots,n))}$ is unbounded), the same is true for the first space (i.e., $n^{\frac{1}{2}}/\| X_{1,n} \|^{(2,3,\cdots,n)}$ is unbounded.)

Proof. Let $C_n = \| X_{1,n} \|^{(2,3,\cdots,n)}$. Then $C_2 = \| 1, 1 \|^{(2)} \leqq 2$, $C_3 = \| \| 1, 1 \|^{(2)}, 1 \|^{(3)} = \| C_2, 1 \|^{(3)} \leqq \| C_2, 1 \|^{(3)\prime}$. We show by induction that $C_n \leqq \| \cdots \| \| C_2, 1 \|^{(2)\prime}, 1 \|^{(3)\prime}, \cdots, 1 \|^{(n)\prime}$. For suppose that $C_{n-1} \leqq \| \cdots \| C_2, 1 \|^{(2)\prime}, \cdots, 1 \|^{(n-1)\prime}$. Then since $C_n = \| C_{n-1}, 1 \|^{(n)} \leqq \| C_{n-1}, 1 \|^{(n)\prime}$, this hypothesis and Lemma 4.1 imply the desired inequality for C_n. Also since $\| x, cy \|^{((i))} = \| x, y \|^{(i)\prime}$ and $\| 1, 0 \|^{((i))} = 1$,

$$\begin{aligned} C_n \leqq \| \cdots \| C_2, 1 \|^{(2)\prime}, \cdots, 1 \|^{(n)\prime} &= \| \cdots \| C_2, c \|^{((2))}, \cdots c \|^{((n))} \\ &\leqq c \| X_{1,n} \|^{((2,3,\cdots,n))} + | C_2 - c | \end{aligned}$$

by the triangle property of the norm. Therefore

$$\frac{n^{\frac{1}{2}}}{C_n} = \frac{n^{\frac{1}{2}}}{\|X_{1,n}\|^{(2,3,\cdots,n)}} \geqq \frac{n^{\frac{1}{2}}}{c\,\|X_{1,n}\|^{((2,3,\cdots,n))} + |C_2 - c|},$$

and this implies the theorem.

THEOREM 5.41. (*Localization Theorem.*) *Suppose we are given a space defined by* $\|x, y\|^{(2)}, \|x, y\|^{(3)}, \cdots$, *where* $\|1, 0\|^{(i)} = \|0, 1\|^{(i)} = 1$ *for all* i, *and a second space defined by* $\|x, y\|^{((2))}, \|x, y\|^{((3))}, \cdots$, *where also* $\|1, 0\|^{((i))} = \|0, 1\|^{((i))} = 1$ *for all* i. *Suppose further that there exists an* $\epsilon > 0$ *such that for each* i

$$\|1, x\|^{(i)} = \|1, x\|^{((i))} \quad \text{if} \quad |x| < \epsilon;$$

i.e., for each i *the corresponding two-dimensional norms coincide* (*uniformly*) *in any neighborhood, however small, of* (1, 0). *Then if by the method of Theorem* 4.3 *it is possible to construct a subspace in one of the spaces on which there exists no projection, it is likewise possible in the other space, no matter how the two sets of norms may differ outside of the* ϵ-*neighborhoods of* (1, 0).

Proof. This theorem follows immediately from Theorem 5.4, by taking $c = \epsilon^{-1}$.

By Theorems 5.4 and 5.41 we are able to obtain in particular obvious generalizations of the result of Theorem 5.2 for $p_2p_3p_4 \cdots$ spaces.

If in a $p_2p_3p_4 \cdots$ space all of the p_i's are equal, $p = p_2 = p_3 = \cdots$, the space is of course l_p. A space for which the defining sequence of two-dimensional norms consists of a single norm taken repeatedly will be called a *repeated* space. By Theorem 5.41 in particular if the norms which define two repeated spaces coincide in a neighborhood of (1, 0), then if in one space it is possible by the method of Theorem 4.3 to construct a subspace on which there exists no projection, the same is true in the other repeated space.

For a repeated space for which the defining two-dimensional norm $\|x, y\|$ has the properties

(1) $\|1, 0\| = \|0, 1\| = 1$,

(2) $\|x, y\|_2 \leqq \|x, y\|$ for all x, y, where $\|x, y\|_2 = (x^2 + y^2)^{\frac{1}{2}}$,

(3) $\|x, y\| = \|y, x\|$ for all x, y

it may be shown that a n.s.c. that the space be isomorphic to Hilbert space is that there exist an $a > 0$ such that $\|x, y\| \leqq \|ax, y\|_2$ for all x, y; i.e., that the curvature (if it is defined) of the curve $\|x, y\| = 1$ at (1, 0) be finite (and a similar condition if (2) is replaced by (2′): $\|x, y\|_2 \geqq \|x, y\|$). Thus if any two norms with properties (1), (2), (3) coincide near (1, 0), then if one of the corresponding repeated spaces is isomorphic to Hilbert space, the same is true of the other; and the Hilbert space character of such a repeated space, like the non-existence of projections (by our method), depends only on the behavior near (1, 0) of the two-dimensional norm which defines the space.

6. **The spaces of Orlicz.** An l_p-space consists of all sequences $x = \{x_i\}$ such that $\sum_1^\infty M(x_i) < \infty$, where $M(u) = |u|^p$. For the Orlicz spaces the function $|u|^p$ is replaced by a more general monotone function.

An *N-function* (Orlicz [8]) is any convex, continuous, monotone increasing function $M(u)$ defined for $0 \leqq u < \infty$ such that $M(0) = 0$. A n.s.c. that the class of sequences $x = \{x_i\}$ such that $\sum_{i=1}^\infty M(|x_i|) < \infty$ be a linear space is that the N-function $M(u)$ have the property

$$(\Delta_2)' \qquad M[2u] \leqq hM[u] \quad \text{for} \quad 0 \leqq u < u_0 \qquad (h > 0).$$

By means of N-functions with additional restrictions, Orlicz defines Banach spaces of sequences and functions in which the norm has certain desired properties. We consider here only spaces of sequences. Instead of Orlicz's norm, we introduce a norm which is analogous to the norm of Theorem 5.3, as follows: The unit sphere is the set of all sequences $x = \{x_i\}$ such that $\sum_1^\infty M(|x_i|) = 1$; and the norm is defined for all sequences of the space by the homogeneity norm property, i.e., for any sequence $x = \{x_i\}$,

$$\|x\| = \frac{1}{k}, \qquad \text{where } \sum_1^\infty M(k|x_i|) = 1.$$

We verify that the unit sphere is convex, so that we do in fact have a norm: If $\{x_i\}$ and $\{y_i\}$ are any two sequences of the unit sphere, by the convexity of of $M(u)$, for each i

$$M\left(\frac{|x_i + y_i|}{2}\right) \leqq \frac{M(|x_i|) + M(|y_i|)}{2},$$

so that $\sum_{i=1}^\infty M(\frac{1}{2}|x_i + y_i|) \leqq 1$; convexity of the unit sphere follows since $M(u)$ is monotone. By §4, our norm is isomorphic to Orlicz's norm whenever Orlicz's norm is defined.

THEOREM 6.1. *Given any Orlicz space of sequences defined by an N-function $M(u)$ with property $(\Delta_2)'$. If there exists a projection on every closed linear subspace, the space must be isomorphic to Hilbert space.*

Proof. The Orlicz space is a symmetric sequence space. Suppose a projection exists on every subspace. If in any subspace $l(1, 2, \cdots, n)$ we introduce a Euclidean norm by $\|x\|_2 = (\sum_1^n |x_i|^2)^{\frac{1}{2}}$, then by Theorem 4.3 the Euclidean radius of the unit sphere in the $(1, 1, \cdots, 1)$ direction must be bounded from 0

and from ∞ as $n \to \infty$. Let a_n be such that $nM(a_n) = 1$. Then the radius is $n^{\frac{1}{2}}a_n$, and we have

$$n^{\frac{1}{2}}a_n \leqq K, \qquad a_n \leqq \frac{K}{n^{\frac{1}{2}}}, \quad \frac{1}{n} = M(a_n) \leqq M\left(\frac{K}{n^{\frac{1}{2}}}\right).$$

For any u such that

$$\frac{1}{(n+1)^{\frac{1}{2}}} \leqq u \leqq \frac{1}{n^{\frac{1}{2}}}, \qquad \frac{n}{n+1}u^2 \leqq \frac{1}{n+1} \leqq M(Ku);$$

hence there exists a $k > 0$ and u_0, such that $ku^2 \leqq M(u)$ for $0 \leqq u < u_0$. We also have $0 < \epsilon < n^{\frac{1}{2}}a_n$ for all n. This implies similarly that there exist a $k_1 > 0$ and u_1 such that $k_1u^2 \geqq M(u)$ for $0 \leqq u < u_1$. We now use the following theorem of Birnbaum and Orlicz: Suppose $M(u)$ and $N(u)$ are any two N-functions. In order that every sequence $x = \{x_i\}$ which is convergent with $M(u)$ be also convergent with $N(u)$, it is necessary and sufficient that there exist numbers $a > 0$ and $b > 0$ such that

$$N(u) \leqq b \cdot M(u) \qquad \text{for } 0 \leqq u \leqq a$$

([4], Satz 5a, p. 5). Therefore any Orlicz space in which projections exist contains the same sequences as Hilbert space l_2, and Theorem 6.1 follows by Theorem 4.1.

Bibliography

1. O. B. Ader, *An affine invariant of convex regions*, this Journal, vol. 4(1938), p. 291.
2. S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.
3. F. Bohnenblust, *Convex regions and projections in Minkowski spaces*, Annals of Math., vol. 39(1938), p. 301.
4. Z. W. Birnbaum and W. Orlicz, *Über die Verallgemeinerung des Begriffes der zueinander konjugierten Potenzen*, Studia Mathematica, vol. 3(1931), pp. 1–67.
5. P. Jordan and J. von Neumann, *On inner products in linear metric spaces*, Annals of Mathematics, vol. 36(1935), p. 719.
6. F. J. Murray, *On complementary manifolds and projections in spaces L_p and l_p*, Transactions of the American Mathematical Society, vol. 41(1937), p. 138.
7. W. Orlicz, *Über konjugierte Exponentenfolgen*, Studia Mathematica, vol. 3(1931), pp. 200–211.
8. W. Orlicz, *Über eine gewisse Klasse von Raumen vom Typus B*, Bulletin International, Académie Polonaise des Sciences, A, vols. 8–9(1932), pp. 207–220.
9. W. Orlicz, *Über Räume* (L^M), Bulletin International, Académie Polonaise des Sciences, A, 1936, pp. 93–107.

Princeton University and Oregon State College.

THE FREDHOLM THEORY OF INTEGRAL EQUATIONS

By F. Smithies

1. Introduction

1.1. Let us consider Fredholm's integral equation of the second kind:

$$x(s) = y(s) + \lambda \int k(s, t)\, x(t)\, dt, \tag{1.1.1}$$

where the integration is taken over a fixed interval (finite or infinite), and the range of the variable s is the same interval. If $k(s, t)$ is continuous, and the interval of integration is finite, the solution is given by Fredholm's famous formulas (see [3][1]):

$$x(s) = y(s) + \frac{\lambda}{d(\lambda)} \int d(s, t; \lambda)\, y(t)\, dt. \tag{1.1.2}$$

To describe the symbols appearing in (1.1.2) and in similar formulas we use the following notation. Let M_n be the determinant whose elements are $k(u_i, u_j)$ $(i, j = 1, 2, \cdots, n)$. Let N_n be the determinant obtained by replacing the elements on the main diagonal of M_n by zero. Let $M_n^*(s, t)$ be the determinant obtained by bordering M_n thus

$$\left| \begin{array}{c|ccc} k(s, t) & k(s, u_1) & \cdots & k(s, u_n) \\ \hline k(u_1, t) & & & \\ \cdots & & M_n & \\ k(u_n, t) & & & \end{array} \right|$$

and let N_n^* be obtained similarly from N_n. In this notation we have

$$d(\lambda) = 1 - \lambda \int k(u, u)\, du + \frac{\lambda^2}{2!} \iint M_2\, du_1\, du_2 - \cdots,$$

$$d(s, t; \lambda) = k(s, t) - \lambda \int M_1^*(s, t)\, du_1 + \frac{\lambda^2}{2!} \iint M_2^*(s, t)\, du_1\, du_2 - \cdots,$$

provided that λ is not a zero of $d(\lambda)$. The series are convergent for all finite complex values of λ.

In the present paper I wish to discuss the solution of the equation (1.1.1) when all we know about $k(s, t)$ is that it is a measurable function of (s, t) and that

$$\iint |k(s, t)|^2\, ds\, dt < \infty.$$

Received July 5, 1940.

[1] Numbers in square brackets refer to the bibliography at the end of the paper.

The interval of integration may now be finite or infinite. In this case the diagonal terms $k(u_\nu, u_\nu)$ need not even be measurable functions, so some modification of the formulas will be necessary. It was shown by Carleman [1] in 1921 that if these diagonal terms are replaced by zeros in $d(\lambda)$ and in $d(s, t; \lambda)$, the solution is given by the formula (1.1.2) so modified, the series still being convergent for all complex values of λ. His proof uses some rather difficult inequalities, and assumes a considerable part of the theory for continuous kernels $k(s, t)$.

In 1928 Hille and Tamarkin ([7], [8]) showed that the integral equation can be transformed into a system of linear equations in an infinity of unknowns, whose solution can then be expressed in terms of infinite determinants. This solution can then be translated into Carleman's form.

In this paper I shall discuss another approach to the problem. Carleman approximates to the given kernel by a continuous one, and Hille and Tamarkin represent the given equation by a system of linear equations in an infinity of unknowns; the method followed here will be to approximate the given equation by a finite system of linear equations in a finite number of unknowns, or, in other words, to approximate the given kernel by a "degenerate" kernel. The solution obtained is equivalent to that given by the modified Fredholm formulas, but it is expressed in terms of the iterates of the kernel and their traces. Such expressions are to be found in the literature, but very little attention seems to have been paid to them, in spite of the fact that they give the only form of the solution that transforms in an obvious and trivial way under an arbitrary unitary transformation of the Hilbert space of functions of integrable square. A formula of this kind is given in the Encyklopädie article of Hellinger and Toeplitz [5], and another is hinted at there. A solution in such terms is given in a paper by Michal and Martin [10], but in a setting so abstract that it is impossible to obtain really powerful results; for instance, they are unable to show that the series obtained are convergent for all values of the parameter λ.

1.2. The main result of the present paper is contained in Theorem 5.5, where the solution in terms of the iterated kernels and their traces is given. In the following theorems the statements about the convergence of the series are progressively strengthened, and in Theorem 5.8 it is shown that this solution can be translated into the form given by Carleman; we therefore have a new proof of Carleman's results.

We shall find it convenient to use the notation and terminology of the theory of linear transformations in Hilbert space (cf. Stone [11] and Julia [9]); in Chapter 2 we shall prove a number of preparatory lemmas in this theory, most of which do not appear to be given in the existing literature. Chapter 3 deals with the Fredholm theory for a finite-dimensional unitary space; in Chapter 4 the modified formulas are introduced, and some necessary inequalities are proved, and in Chapter 5 the results are extended to the general case, and the main theorems mentioned above are reached.

We have confined ourselves to proving that the solution is given by our formulas when the modified Fredholm determinant $\delta(\lambda)$ does not vanish, and have not gone into the question of solutions of the homogeneous equation. The well-known results can in fact be obtained by easy modifications of the familiar methods.

2. Preparatory lemmas

2.1. We shall be considering bounded linear transformations defined throughout a Hilbert space $\mathfrak{H}$ (Stone [11], p. 3) or a finite-dimensional unitary space $\mathfrak{M}$ (Stone [11], p. 16). We shall in general use Stone's notation and terminology, except that we write $\| x \| = (x, x)^{\frac{1}{2}}$. The treatment as given will apply to complex spaces, but the results obtained hold also for real Hilbert space and for finite-dimensional Euclidean spaces, the necessary changes in the proofs being quite insignificant.

It is important to remark that every finite-dimensional linear manifold of $\mathfrak{H}$ is closed, and is therefore a finite-dimensional unitary space. (Cf. Stone [11], p. 19.)

2.2. If A is a bounded linear transformation, and I is the identical transformation,[2] we can form polynomials $\lambda_0 I + \lambda_1 A + \cdots + \lambda_n A^n$ with complex coefficients; these can be added, subtracted and multiplied like ordinary polynomials, and can be multiplied by complex numbers. In particular, we can form determinants involving these polynomials, and we shall have frequent occasion to do so.

We shall denote the *bound* of A, i.e., the least number λ such that $\| Ax \| \leqq \lambda \| x \|$ for all $x \in \mathfrak{H}$, by $| A |$. We then have

$$| \mu A | = | \mu | \cdot | A |, \qquad | A + B | \leqq | A | + | B |, \qquad | AB | \leqq | A | \cdot | B |,$$

for arbitrary transformations[3] A, B and arbitrary complex μ.

If $\{A_n\}$ is a sequence of transformations such that

$$| A_n - A | \to 0 \qquad (n \to \infty),$$

we say that $\{A_n\}$ is *uniformly convergent* to A, and that $A_n \to A$ uniformly as $n \to \infty$. It follows from the completeness of Hilbert space that, if $| A_m - A_n | \to 0$ as $m, n \to \infty$, then there is a transformation A such that $| A_n - A | \to 0$ as $n \to \infty$. If $A_n \to A$ uniformly, and B is an arbitrary transformation, then $BA_n \to BA$ uniformly, and $A_n B \to AB$ uniformly.

2.3. If the value of (Ax, y) is given for all $x, y \in \mathfrak{H}$, or even for all the elements of a complete orthonormal set $\{x_\alpha\}$, then the transformation A is completely determined (cf. Stone [11], pp. 63, 88). The *adjoint* transformation A^* of A is defined by the relation $(Ax, y) = (x, A^*y)$, holding for all $x, y \in \mathfrak{H}$. A^* is

[2] We *define* $A^0 = I$; this convention is consistent with the ordinary laws of algebra.

[3] We shall usually omit the words "bounded linear", since we shall not have occasion to consider any other kind of transformation.

again a bounded linear transformation defined throughout $\mathfrak{H}$, and $A^{**} = A$. Since

$$| A | = \sup_{||x||\cdot||y|| \leqq 1} | (Ax, y) |,$$

we have $| A^* | = | A |$.

2.4. If $\{x_\alpha\}$ is a complete orthonormal set, we define the *matrix* of A with respect to it as $[a_{\alpha\beta}]$, where

$$a_{\alpha\beta} = (Ax_\beta , x_\alpha) \qquad (\alpha, \beta = 1, 2, \cdots).$$

If A has the matrix $[a_{\alpha\beta}]$ and B has the matrix $[b_{\alpha\beta}]$, then the matrix of λA is $[\lambda a_{\alpha\beta}]$, that of $A + B$ is $[a_{\alpha\beta} + b_{\alpha\beta}]$, and that of AB is

$$[\sum_{\gamma=1}^{\infty} a_{\alpha\gamma} b_{\gamma\beta}];$$

the matrix of A^* is $[c_{\alpha\beta}]$, where $c_{\alpha\beta} = \bar{a}_{\beta\alpha}$.

If $(x, x_\alpha) = \xi_\alpha$ and $(y, x_\alpha) = \eta_\alpha$ $(\alpha = 1, 2, \cdots)$, we have

$$(Ax, y) = \sum_{\alpha,\beta=1}^{\infty} a_{\alpha\beta}\xi_\beta\bar{\eta}_\alpha .$$

2.5. If A has the matrix $[a_{\alpha\beta}]$, and

$$|| A ||^2 = \sum_{\alpha,\beta=1}^{\infty} | a_{\alpha\beta} |^2 < \infty,$$

A is said to be of *finite norm*. $|| A ||$ is called the *norm* of A, and is independent of the particular orthonormal set used in its definition. (Cf. Stone [11], p. 66.)

We have $|| \mu A || = | \mu |\cdot|| A ||$, $|| A^* || = || A ||$,

$$|| A + B || \leqq || A || + || B ||, \qquad || AB || \leqq || A ||\cdot|| B ||.$$

In particular, $|| A^n || \leqq || A ||^n$ $(n = 1, 2, \cdots)$. If $\{A_n\}$ is a sequence of transformations of finite norm such that $|| A_n - A || \to 0$ as $n \to \infty$, then A is of finite norm, and we say that $\{A_n\}$ is *convergent in norm* to A and that $A_n \to A$ *in norm* as $n \to \infty$. If $\{A_n\}$ is a sequence of transformations of finite norm such that $|| A_m - A_n || \to 0$ as $m, n \to \infty$, there is a transformation A of finite norm such that $|| A_n - A || \to 0$ as $n \to \infty$.

Since, by repeated application of Cauchy's inequality, we have $| (Ax, y) | \leqq || A ||\cdot|| x ||\cdot|| y ||$, it follows that $| A | \leqq || A ||$, and that convergence in norm implies uniform convergence; the converse statement is false.

2.6. Lemma. *If A is an arbitrary transformation and K is of finite norm, then AK and KA are of finite norm, and $|| AK || \leqq | A |\cdot|| K ||$, $|| KA || \leqq | A |\cdot|| K ||$.*

We write $a_{\alpha\beta} = (Ax_\beta, x_\alpha)$, $k_{\alpha\beta} = (Kx_\beta, x_\alpha)$. Then the matrix of AK is

$$[\sum_{\gamma=1}^{\infty} a_{\alpha\gamma} k_{\gamma\beta}].$$

By the definition of $| A |$,

$$\sum_{\alpha=1}^{\infty} | \sum_{\gamma=1}^{\infty} a_{\alpha\gamma} k_{\gamma\beta} |^2 \leqq | A |^2 \sum_{\gamma=1}^{\infty} | k_{\gamma\beta} |^2 \qquad (\beta = 1, 2, \cdots).$$

Hence

$$\sum_{\alpha,\beta=1}^{\infty} | \sum_{\gamma=1}^{\infty} a_{\alpha\gamma} k_{\gamma\beta} |^2 \leqq | A |^2 \sum_{\beta,\gamma=1}^{\infty} | k_{\gamma\beta} |^2;$$

i.e.,

$$\| AK \|^2 \leqq | A |^2 \cdot \| K \|^2, \qquad \| AK \| \leqq | A | \cdot \| K \|.$$

Finally,

$$\| KA \| = \| (KA)^* \| = \| A^*K^* \| \leqq | A^* | \cdot \| K^* \| = | A | \cdot \| K \|.$$

COROLLARY. *If $K_n \to K$ in norm, and A is arbitrary, then $AK_n \to AK$ in norm, and $K_nA \to KA$ in norm.*

2.7. If $A = BC$, where B and C are of finite norm, we define the *trace* of A as

$$\tau(A) = \sum_{\alpha=1}^{\infty} a_{\alpha\alpha} = \sum_{\alpha,\beta=1}^{\infty} b_{\alpha\beta} c_{\beta\alpha}.$$

Since

$$| \tau(A) |^2 \leqq \sum_{\alpha,\beta=1}^{\infty} | b_{\alpha\beta} |^2 \sum_{\alpha,\beta=1}^{\infty} | c_{\beta\alpha} |^2,$$

$\tau(A)$ is finite, and $| \tau(A) | \leqq \| B \| \cdot \| C \|$. The trace of A depends neither on the way A is resolved into a product BC nor on the orthonormal system used in the definition.

The definition is extended to transformations that can be expressed as the sum of a finite number of such products by means of the equation

$$\tau(A_1 + A_2) = \tau(A_1) + \tau(A_2).$$

We have

$$\tau(BC) = \sum_{\alpha,\beta=1}^{\infty} b_{\alpha\beta} c_{\beta\alpha} = \tau(CB), \qquad \tau(A^*) = \overline{\tau(A)}.$$

If $A_n \to A$ in norm and $B_n \to B$ in norm, then

$$\begin{aligned} &| \tau(A_nB_n) - \tau(AB) | \\ &\quad = | \tau[(A_n - A)(B_n - B)] + \tau[A(B_n - B)] + \tau[(A_n - A)B] | \\ &\quad \leqq \| A_n - A \| \cdot \| B_n - B \| + \| A \| \cdot \| B_n - B \| + \| A_n - A \| \cdot \| B \| \to 0; \end{aligned}$$

i.e., $\tau(A_n B_n) \to \tau(AB)$. If A is of finite norm, then $\tau(A^n)$ exists when $n \geqq 2$, and

$$| \tau(A^n) | = | \tau(A^{n-1} \cdot A) | \leqq \| A^{n-1} \| \cdot \| A \| \leqq \| A \|^n.$$

2.8. If E is any space in which a Lebesgue measure and a Lebesgue integral can be defined, it is well known that the complex-valued measurable functions $x(t)$, defined in E and such that

$$\int_E | x(t) |^2 \, dt < \infty,$$

form a Hilbert space.[4] In particular, this is so if E is a measurable subset of positive measure of p-dimensional Euclidean space and the measure function is ordinary p-dimensional Lebesgue measure. We denote the Hilbert space by $\mathfrak{L}^2(E)$ or $\mathfrak{L}^2$. If $x \in \mathfrak{L}^2$, $y \in \mathfrak{L}^2$, then

(i) $x = y$ if and only if $x(t) = y(t)$ for almost all t,

(ii) $x + y$ corresponds to $x(t) + y(t)$,

(iii) λx corresponds to $\lambda x(t)$, and[5]

(iv) $(x, y) = \int x(t)\overline{y(t)} \, dt$.

The transformation $y = Kx$ defined by[6]

$$y(s) = \int k(s, t) x(t) \, dt,$$

where $k(s, t)$ is a measurable function in the space $E \times E$ such that

$$\iint_{E \times E} | k(s, t) |^2 \, ds \, dt < \infty,$$

is a transformation of finite norm in $\mathfrak{L}^2$, and

$$\| K \|^2 = \iint | k(s, t) |^2 \, ds \, dt.$$

Conversely, if K is a transformation of finite norm in $\mathfrak{L}^2$, then there is a function $k(s, t)$ determining the transformation in the way just described. The function $k(s, t)$ is called the *kernel* of K. If K and L have the kernels $k(s, t)$ and $l(s, t)$ respectively, then

(i) $K = L$ if and only if $k(s, t) = l(s, t)$ for almost all (s, t),

(ii) $K + L$ has the kernel $k(s, t) + l(s, t)$,

(iii) λK has the kernel $\lambda k(s, t)$,

(iv) K^* has the kernel $\overline{k(t, s)}$,

(v) KL has the kernel $\int k(s, u) l(u, t) \, du$.

[4] In certain very special cases, they form a finite-dimensional unitary space.

[5] We omit the range of integration when no ambiguity can arise.

[6] Equalities and inequalities between functions will in general be understood to hold almost everywhere.

We denote the kernel of K^n by $k_n(s, t)$. The trace of KL is given by

$$\tau(KL) = \iint k(s, t)l(t, s)\, ds\, dt.$$

2.9. LEMMA. *Let A be a bounded transformation in $\mathfrak{L}^2$, and let K and L be transformations of finite norm with kernels $k(s, t)$ and $l(s, t)$ respectively. Write*

$$a(s) = \left\{\int |\, k(s, t)\,|^2\, dt\right\}^{\frac{1}{2}}, \qquad b(t) = \left\{\int |\, l(s, t)\,|^2\, ds\right\}^{\frac{1}{2}}.$$

Let $KAL = G$, and let $g(s, t)$ be the kernel of G.[7] *Then*

$$|\, g(s, t)\,| \leqq |\, A\,| \cdot a(s)b(t)$$

for almost all (s, t).

For almost all t, $l(s, t)$, regarded as a function of s, defines an element l_t of $\mathfrak{L}^2$. Let $x \in \mathfrak{L}^2$, $y \in \mathfrak{L}^2$, and write $A^*x = z$, $Al_t = p_t$, $p_t(s) = p(s, t)$. Then

$$\begin{aligned}(ALy, x) &= (Ly, A^*x) = (Ly, z)\\ &= \iint l(s, t)\overline{z(s)}y(t)\, ds\, dt\\ &= \int y(t)\, dt \int l_t(s)\overline{z(s)}\, ds.\end{aligned}$$

Now

$$\begin{aligned}\int l_t(s)\overline{z(s)}\, ds &= (l_t, z)(l_t, A^*x) = (Al_t, x)\\ &= (p_t, x) = \int p_t(s)\overline{x(s)}\, ds = \int p(s, t)\overline{x(s)}\, ds.\end{aligned}$$

Hence

$$\begin{aligned}(ALy, x) &= \int y(t)\, dt \int p(s, t)\overline{x(s)}\, ds\\ &= \iint p(s, t)\overline{x(s)}y(t)\, ds\, dt;\end{aligned}$$

consequently $p(s, t)$ is the kernel of AL. Since, by Lemma 2.6, AL is of finite norm, we have

$$\iint |\, p(s, t)\,|^2\, ds\, dt < \infty.$$

[7] The transformation G is of finite norm, by Lemma 2.6.

Now write $K^*x = w$. Then

$$(KALy, x) = (ALy, K^*x) = (ALy, w)$$

$$= \iint p(u, t)\overline{w(u)}y(t)\, du\, dt = \iint p(u, t)y(t)\, du\, dt \int k(s, u)\overline{x(s)}\, ds$$

$$= \iint \overline{x(s)}y(t)\, ds\, dt \int k(s, u)p(u, t)\, du.$$

Thus the kernel of $G = KAL$ is

$$g(s, t) = \int k(s, u)p(u, t)\, du.$$

Hence, for almost all (s, t),

$$| g(s, t) | \leqq \left\{\int | k(s, u) |^2 du\right\}^{\frac{1}{2}} \left\{\int | p(u, t) |^2 du\right\}^{\frac{1}{2}}$$

$$= a(s)\cdot\| p_t \| \leqq a(s)\cdot| A |\cdot\| l_t \|$$

$$= a(s)\cdot| A |\cdot\left\{\int | l(s, t) |^2 ds\right\}^{\frac{1}{2}} = | A |\cdot a(s)b(t),$$

the required result.

3. Transformations in a finite-dimensional space

3.1. Let $\mathfrak{M}$ be a p-dimensional unitary space, and let $[x_1, x_2, \cdots, x_p]$ be a complete orthonormal set of vectors in $\mathfrak{M}$; we keep it fixed throughout this chapter. Since $\mathfrak{M}$ is finite-dimensional, every linear transformation K is bounded and of finite norm.

If K has the matrix $[k_{\alpha\beta}]$ $(\alpha, \beta = 1, \cdots, p)$, and K^n has the matrix $[k_{\alpha\beta}^{(n)}]$ $(n > 1)$, we write

$$\sigma_1 = \tau(K) = \sum_{\alpha=1}^{p} k_{\alpha\alpha}, \qquad \sigma_n = \tau(K^n) = \sum_{\alpha=1}^{p} k_{\alpha\alpha}^{(n)}.$$

We may regard σ_1 as being $\tau(IK)$, if we wish to conform with the treatment given in §2.7; the identical transformation I is of finite norm in $\mathfrak{M}$.

We also write $\det K = \det [k_{\alpha\beta}]$. The polynomial $\varphi(\kappa) = \det (\kappa I - K)$ is called the *characteristic polynomial* of K; we denote its zeros by $\kappa_1, \kappa_2, \cdots, \kappa_p$, repeating them according to their multiplicity.

The numbers $\sigma_1, \sigma_2, \cdots, \det K, \kappa_1, \kappa_2, \cdots, \kappa_p$, and the polynomial $\varphi(\kappa)$ are all independent of the particular orthonormal system $[x_1, x_2, \cdots, x_p]$ used in defining them.

3.2. Lemma. *If $g(t)$ is an arbitrary polynomial with complex coefficients, then the zeros (repeated according to their multiplicity) of* $\det(\kappa I - g(K))$ *are* $g(\kappa_1)$ $g(\kappa_2), \cdots, g(\kappa_p)$.

For a proof, see, e.g., Courant and Hilbert [2], p. 19.

3.3. Lemma.

$$\sigma_n = \sum_{\alpha=1}^{p} \kappa_\alpha^n \qquad (n = 1, 2, \cdots).$$

We have

$$\varphi(\kappa) = \prod_{\alpha=1}^{p} (\kappa - \kappa_\alpha) = \det(\kappa I - K). \tag{3.3.1}$$

Equating the coefficients of κ^{p-1} in (3.3.1), we obtain

$$\sum_{\alpha=1}^{p} \kappa_\alpha = \sum_{\alpha=1}^{p} k_{\alpha\alpha} = \sigma_1.$$

If we apply this result to K^n instead of K, and use Lemma 3.2, the required result follows.

3.4. We now consider the equation

$$(I - \lambda K)x = y, \tag{3.4.1}$$

where K, y and λ are given, and x is to be determined. If

$$x = \sum_{\alpha=1}^{p} \xi_\alpha x_\alpha, \qquad y = \sum_{\alpha=1}^{p} \eta_\alpha x_\alpha,$$

(3.4.1) is equivalent to

$$\xi_\alpha = \eta_\alpha + \lambda \sum_{\beta=1}^{p} k_{\alpha\beta}\xi_\beta \qquad (\alpha = 1, 2, \cdots, p), \tag{3.4.2}$$

If λ is not the reciprocal of a zero of $\varphi(\kappa)$, (3.4.2) has a unique solution which can be written in the form

$$\xi_\alpha = [d(\lambda)]^{-1} \sum_{\beta=1}^{p} d_{\alpha\beta}(\lambda)\eta_\beta \qquad (\alpha = 1, 2, \cdots, p), \tag{3.4.3}$$

where $d(\lambda) = \det(I - \lambda K)$. When $\lambda \neq 0$, $d(\lambda) = \lambda^p\varphi(\lambda^{-1})$, and $d(0) = 1$. Equation (3.4.3) may be written still more concisely as

$$x = [d(\lambda)]^{-1}D(\lambda)y,$$

where $D(\lambda)$ is, for each value of λ, a linear transformation in $\mathfrak{M}$.

It can easily be verified that, for arbitrary $x, y \in \mathfrak{M}$,

$$(D(\lambda)y, x) = -\begin{vmatrix} 0 & \xi_1 \cdots \xi_p \\ \eta_1 & \\ \cdots & I - \lambda K \\ \eta_p & \end{vmatrix}. \tag{3.4.4}$$

3.5. These facts may be summed up as follows.

THEOREM. *If* $d(\lambda) \neq 0$, *the equation*

$$(I - \lambda K)x = y$$

has the unique solution

$$x = [d(\lambda)]^{-1}D(\lambda)y,$$

where $D(\lambda)$ *is defined by* (3.4.5), *and*

$$d(\lambda) = \det (I - \lambda K).$$

3.6. Both $d(\lambda)$ and $D(\lambda)$ are polynomials in λ; we may write them formally as power series:

$$d(\lambda) = \sum_{n=0}^{\infty} d_n \lambda^n, \qquad D(\lambda) = \sum_{n=0}^{\infty} D_n \lambda^n, \tag{3.6.1}$$

where the coefficients D_n are linear transformations in $\mathfrak{M}$. Our next task is to obtain explicit expressions for d_n and D_n.

3.7. THEOREM.

$$d_0 = 1, \qquad d_n = \frac{(-1)^n}{n!} P_n \qquad (n = 1, 2, \cdots),$$

where

$$P_n = \begin{vmatrix} \sigma_1 & n-1 & 0 & \cdots & 0 & 0 \\ \sigma_2 & \sigma_1 & n-2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_{n-1} & \sigma_{n-2} & \sigma_{n-3} & \cdots & \sigma_1 & 1 \\ \sigma_n & \sigma_{n-1} & \sigma_{n-2} & \cdots & \sigma_2 & \sigma_1 \end{vmatrix}.$$

We have

$$d(\lambda) = \prod_{\alpha=1}^{p} (1 - \kappa_\alpha \lambda).$$

Hence, if $|\lambda|$ is sufficiently small,

$$\frac{d'(\lambda)}{d(\lambda)} = -\sum_{\alpha=1}^{p} \frac{\kappa_\alpha}{1 - \kappa_\alpha \lambda}$$

$$= -\sum_{\alpha=1}^{p} \sum_{n=0}^{\infty} \kappa_\alpha^{n+1} \lambda^n.$$

Now, by Lemma 3.3,

$$\sum_{\alpha=1}^{p} \kappa_\alpha^{n+1} = \sigma_{n+1}.$$

Consequently

$$\frac{d'(\lambda)}{d(\lambda)} = -\sum_{n=0}^{\infty} \sigma_{n+1}\lambda^n. \tag{3.7.1}$$

We also have

$$d(\lambda) = \sum_{n=0}^{\infty} d_n\lambda^n,$$

so that

$$d'(\lambda) = \sum_{n=0}^{\infty} (n+1)\, d_{n+1}\lambda^n.$$

Combining the last three equations, we obtain

$$\sum_{n=0}^{\infty} (n+1)d_{n+1}\lambda^n = -\sum_{m=0}^{\infty} d_m\lambda^m \sum_{q=0}^{\infty} \sigma_{q+1}\lambda^q. \tag{3.7.2}$$

Equate the coefficients of λ^n on the two sides of (3.7.2); this gives

$$(n+1)\, d_{n+1} = -\sum_{m+q=n} d_m\sigma_{q+1} \qquad (n = 0, 1, \cdots). \tag{3.7.3}$$

Solving the first n of these equations shows the truth of Theorem 3.7.

3.8. Theorem.

$$D_n = \frac{(-1)^n}{n!}\left|\begin{array}{c|cccc} I & n & 0 & \cdots & 0 \\ \hline K & & & & \\ K^2 & & P_n & & \\ \cdots & & & & \\ K^n & & & & \end{array}\right| \qquad (n = 0, 1, 2, \cdots). \tag{3.8.1}$$

To prove this we go back to the fact that, when $d(\lambda) \neq 0$, the unique solution of the equation

$$(I - \lambda K)x = y$$

is given by $x = [d(\lambda)]^{-1}D(\lambda)y$. If we substitute this solution in the equation, we obtain

$$D(\lambda)y = d(\lambda)y + \lambda KD(\lambda)y.$$

This holds for arbitrary $y \in \mathfrak{M}$; hence

$$D(\lambda) = d(\lambda)I + \lambda KD(\lambda).$$

In this equation we replace $d(\lambda)$ and $D(\lambda)$ by the expansions (3.6.1), so obtaining

$$\sum_{n=0}^{\infty} \lambda^n D_n = \sum_{n=0}^{\infty} \lambda^n d_n I + \lambda \sum_{n=0}^{\infty} \lambda^n KD_n. \tag{3.8.2}$$

We recall that these apparently infinite series are in fact finite.

Now equate the coefficients of λ^n on the two sides of (3.8.2); this gives

$$D_0 = I, \qquad D_n = d_n I + K D_{n-1} \qquad (n = 1, 2, \cdots). \tag{3.8.3}$$

Denote the expression on the right side of (3.8.1) by E_n ($n = 0, 1, 2, \cdots$). Then $E_0 = I$; when $n \geqq 1$, we see by expanding the determinant in terms of its first row that

$$E_n = d_n I - n\left(-\frac{1}{n}\right) K E_{n-1} = d_n I + K E_{n-1}.$$

Thus E_n satisfies the recurrence relation (3.8.3), and $E_0 = I = D_0$; hence

$$D_n = E_n \qquad (n = 0, 1, 2, \cdots).$$

This is the required result.

4. Modification of the formulas

4.1. The discussion in Chapter 3 deals satisfactorily with transformations in a finite-dimensional unitary space. The results are not yet, however, in such a form that it is possible to extend them at once to transformations of finite norm in Hilbert space. For an arbitrary transformation K of finite norm, the trace

$$\tau(K) = \sum_{\alpha=1}^{\infty} k_{\alpha\alpha} \tag{4.1.1}$$

does not in general exist; the right side of (4.1.1) is not necessarily convergent. In order to make the extension of the results possible, it will therefore be necessary to make some slight modifications in the formulas of Chapter 3.

4.2. Theorem. *If $d(\lambda) \neq 0$, the equation*

$$(I - \lambda K)x = y$$

has the unique solution

$$x = [\delta(\lambda)]^{-1}\Delta(\lambda)y,$$

where

$$\delta(\lambda) = e^{\sigma_1 \lambda} d(\lambda), \qquad \Delta(\lambda) = e^{\sigma_1 \lambda} D(\lambda).$$

This follows at once from Theorem 3.5.

4.3. The function $\delta(\lambda)$ is an integral function of λ; we can therefore write

$$\delta(\lambda) = \sum_{n=0}^{\infty} \delta_n \lambda^n,$$

the series being convergent for all λ. Similarly, each element of the matrix $\Delta(\lambda)$ can be expanded in an everywhere convergent power series in λ; we express this fact symbolically by writing

$$\Delta(\lambda) = \sum_{n=0}^{\infty} \Delta_n \lambda^n,$$

where, for every value of n, Δ_n is a linear transformation in $\mathfrak{M}$. We must now obtain explicit formulas for δ_n and Δ_n.

4.4. Theorem. $\delta_0 = 1$,

$$\delta_n = \frac{(-1)^n}{n!} Q_n \qquad (n = 1, 2, \cdots),$$

where

$$Q_n = \begin{vmatrix} 0 & n-1 & 0 & \cdots & 0 & 0 & 0 \\ \sigma_2 & 0 & n-2 & \cdots & 0 & 0 & 0 \\ \sigma_3 & \sigma_2 & 0 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_{n-1} & \sigma_{n-2} & \sigma_{n-3} & \cdots & \sigma_2 & 0 & 1 \\ \sigma_n & \sigma_{n-1} & \sigma_{n-2} & \cdots & \sigma_3 & \sigma_2 & 0 \end{vmatrix}.$$

Since $\delta(\lambda) = e^{\sigma_1 \lambda} d(\lambda)$, it follows from equation (3.7.1) that, when $|\lambda|$ is sufficiently small,

$$\frac{\delta'(\lambda)}{\delta(\lambda)} = \sigma_1 + \frac{d'(\lambda)}{d(\lambda)} = -(\sigma_2 \lambda + \sigma_3 \lambda^2 + \cdots). \tag{4.4.1}$$

This differs from (3.7.1) only by the disappearance of the term in σ_1. Hence we can obtain the coefficients δ_n simply by replacing σ_1 by 0 in the formulas for d_n. This gives the required result.

4.5. Theorem. $\Delta_0 = I$,

$$\Delta_n = \frac{(-1)^n}{n!} \left| \begin{array}{c|cccc} I & n & 0 & \cdots & 0 \\ \hline K & & & & \\ K^2 & & & Q_n & \\ \cdots & & & & \\ K^n & & & & \end{array} \right| \qquad (n = 1, 2, \cdots).$$

This is proved by the same argument as Theorem 3.8; the fact that genuinely infinite series are now involved makes no difference.

Corollary. $\tau(K\Delta_{n-1} - \delta_{n-1}K) = -n\delta_n \ (n = 1, 2, \cdots)$.

4.6. We now require a number of inequalities; these will be used when we go over to the Hilbert space case.

LEMMA.[8] *Let $f(z)$ be an integral function of the complex variable $z = re^{i\theta}$, so that*

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

for all z, and suppose that $|f(z)| \leqq g(|z|) = g(r)$ for all z. Then, for all $r \geqq 0$,

$$|a_n| \leqq r^{-n} g(r) \qquad (n = 0, 1, 2, \cdots).$$

4.7. LEMMA.[9] *If A has the matrix $[a_{\alpha\beta}]$, then*

$$|\det A|^2 \leqq \prod_{\alpha=1}^{p} \sum_{\beta=1}^{p} |a_{\alpha\beta}|^2.$$

4.8. LEMMA.

$$|\delta_n| \leqq \frac{e^{\frac{1}{2}n} \|K\|^n}{n^{\frac{1}{2}n}} \qquad (n = 1, 2, \cdots).$$

This holds also when $n = 0$, if we adopt the convention that in this case the right side stands for 1. We may suppose that $\|K\| \neq 0$; for if $\|K\| = 0$, then $\delta_n = 0$ $(n \geqq 1)$, and the lemma is trivial. We have

$$\begin{aligned}\delta(\lambda) &= e^{\sigma_1 \lambda} d(\lambda) \\ &= \exp\,(\lambda \sum_{\alpha=1}^{p} k_{\alpha\alpha}) \det\,(I - \lambda K).\end{aligned}$$

It now follows from Lemma 4.7 that

$$\begin{aligned}|\delta(\lambda)|^2 &\leqq \{|\exp\,(\lambda k_{11})|^2(|1 - \lambda k_{11}|^2 + |\lambda k_{12}|^2 + \cdots + |\lambda k_{1p}|^2)\} \cdots \\ &\qquad \{|\exp\,(\lambda k_{pp})|^2(|\lambda k_{p1}|^2 + |\lambda k_{p2}|^2 + \cdots + |1 - \lambda k_{pp}|^2)\} \\ &= \prod_{\alpha=1}^{p} \{\exp\,[2\Re(\lambda k_{\alpha\alpha})][1 - 2\Re(\lambda k_{\alpha\alpha}) + |\lambda|^2 \sum_{\beta=1}^{p} |k_{\alpha\beta}|^2]\}.\end{aligned}$$

We now use the inequality $1 + a \leqq e^a$, which holds for all real a. This gives us

$$\begin{aligned}|\delta(\lambda)|^2 &\leqq \prod_{\alpha=1}^{p} \exp\,[2\Re(\lambda k_{\alpha\alpha}) - 2\Re(\lambda k_{\alpha\alpha}) + |\lambda|^2 \sum_{\beta=1}^{p} |k_{\alpha\beta}|^2] \\ &= \exp\,[|\lambda|^2 \cdot \|K\|^2];\end{aligned}$$

i.e.,

$$|\delta(\lambda)| \leqq \exp\,[\tfrac{1}{2}|\lambda|^2 \cdot \|K\|^2].$$

We next apply Lemma 4.6 to the integral function $\delta(\lambda)$, obtaining

$$|\delta_n| \leqq r^{-n} \exp\,[\tfrac{1}{2} r^2 \|K\|^2] \qquad (n = 0, 1, 2, \cdots)$$

[8] See, e.g., Titchmarsh [12], p. 84.

[9] This is Hadamard's inequality; see, e.g., Hardy, Littlewood and Pólya [4], pp. 34–36.

for all $r \geqq 0$. Take $r = n^{\frac{1}{2}} \| K \|^{-1}$; then

$$| \delta_n | \leqq \frac{e^{\frac{1}{2}n} \| K \|^n}{n^{\frac{1}{2}n}} \qquad (n = 1, 2, \cdots).$$

This is the required result.

4.9. Lemma.

$$| \Delta_n | \leqq \frac{e^{\frac{1}{2}n+\frac{1}{2}} \| K \|^n}{n^{\frac{1}{2}n}} \qquad (n = 1, 2, \cdots).$$

To make this hold for $n = 0$, the right side must then be taken to be $e^{\frac{1}{2}}$.

We may again suppose that $\| K \| \neq 0$. For arbitrary $x, y \in \mathfrak{M}$, we have, by equation (3.4.5),

$$\begin{aligned}(\Delta(\lambda)y,\ x) &= e^{\sigma_1\lambda}(D(\lambda)y,\ x) \\ &= -\exp\left(\lambda \sum_{\alpha=1}^{p} k_{\alpha\alpha}\right) \begin{vmatrix} 0 & \bar{\xi}_1 \cdots \bar{\xi}_p \\ \eta_1 & \\ \cdots & I - \lambda K \\ \eta_p & \end{vmatrix}.\end{aligned}$$

Take $\| x \| = 1$, $\| y \| = 1$. Then, using Lemma 4.7, and going through the same process as in the proof of Lemma 4.8, we obtain

$$| (\Delta(\lambda)y,\ x) |^2 \leqq e^{1+|\lambda|^2\|K\|^2}.$$

Hence, for arbitrary $x, y \in \mathfrak{M}$,

$$| (\Delta(\lambda)y,\ x) | \leqq e^{\frac{1}{2}+\frac{1}{2}|\lambda|^2\|K\|^2} \| x \| \cdot \| y \|.$$

Now apply Lemma 4.6; we get

$$| (\Delta_n y,\ x) | \leqq \frac{e^{\frac{1}{2}+\frac{1}{2}n} \| K \|^n}{n^{\frac{1}{2}n}} \| x \| \cdot \| y \| \qquad (n = 1, 2, \cdots)$$

for any $x, y \in \mathfrak{M}$; consequently

$$| \Delta_n | \leqq \frac{e^{\frac{1}{2}+\frac{1}{2}n} \| K \|^n}{n^{\frac{1}{2}n}} \qquad (n = 1, 2, \cdots).$$

This is the required result.

5. The formulas in Hilbert space

5.1. We now consider transformations K of finite norm defined in Hilbert space. We choose a complete orthonormal set $[x_1, x_2, \cdots]$, which will be kept fixed throughout; we have

$$\| K \|^2 = \sum_{\alpha,\beta=1}^{\infty} | k_{\alpha\beta} |^2,$$

where

$$k_{\alpha\beta} = (Kx_\beta , x_\alpha) \qquad (\alpha, \beta = 1, 2, \cdots).$$

We recall from §2.7 that $\tau(K^n)$ exists when $n \geqq 2$, and write

$$\tau(K^n) = \sigma_n \qquad (n = 2, 3, \cdots).$$

Then

$$| \sigma_n | \leqq \| K \|^n \qquad (n = 2, 3, \cdots).$$

5.2. If p is any positive integer, we define the transformation E_p by the equation

$$E_p \left(\sum_{\alpha=1}^{\infty} \xi_\alpha x_\alpha \right) = \sum_{\alpha=1}^{p} \xi_\alpha x_\alpha .$$

E_p is defined throughout $\mathfrak{H}$, and $E_p^2 = E_p$. In the usual terminology, E_p is the projection on the closed linear manifold $\mathfrak{M}_p$ determined by $x_1, x_2, \cdots, x_p$.

We write $E_pKE_p = K_p$ $(p = 1, 2, \cdots)$; then

$$(K_p x_\beta, x_\alpha) = \begin{cases} k_{\alpha\beta} & (\alpha, \beta = 1, 2, \cdots, p), \\ 0 & (\alpha > p \text{ or } \beta > p). \end{cases}$$

We also write $\sigma_{n,p} = \tau(K_p^n)$.

We may equally well regard K_p as being a linear transformation in the finite-dimensional unitary space $\mathfrak{M}_p$; the whole theory of Chapters 3 and 4 can then be applied to it. This point of view makes no difference to the values of $\| K_p \|$ or $\sigma_{n,p}$; the only change is that the rôle of the identical transformation I is played by E_p.

5.3. Lemma. *Define K_p $(p = 1, 2, \cdots)$ as in §5.2. Then*

(i) $$\lim_{p\to\infty} \| K_p^n - K^n \| = 0 \qquad (n = 1, 2, \cdots),$$

(ii) $$\lim_{p\to\infty} \| K_p^n \| = \| K^n \| \qquad (n = 1, 2, \cdots),$$

(iii) $$\lim_{p\to\infty} \sigma_{n,p} = \sigma_n \qquad (n = 2, 3, \cdots),$$

(iv) *for arbitrary $x, y \in \mathfrak{H}$,*

$$\lim_{p\to\infty} (K_p^n x, y) = (K^n x, y) \qquad (n = 1, 2, \cdots).$$

We begin by remarking that

$$\| K_p \|^2 = \sum_{\alpha,\beta=1}^{p} | k_{\alpha\beta} |^2 \leqq \sum_{\alpha,\beta=1}^{\infty} | k_{\alpha\beta} |^2 = \| K \|^2,$$

so that $\| K_p^n \| \leqq \| K_p \|^n \leqq \| K \|^n$ for all n; secondly,

$$\| K - K_p \|^2 = \sum_{\alpha,\beta=1}^{\infty} | k_{\alpha\beta} |^2 - \sum_{\alpha,\beta=1}^{p} | k_{\alpha\beta} |^2$$
$$\to 0 \qquad (p \to \infty).$$

We now prove (i) by induction. Suppose that

$$\| K^{n-1} - K_p^{n-1} \| \to 0 \qquad (p \to \infty).$$

Then

$$\begin{aligned} \| K^n - K_p^n \| &\leqq \| K^n - K^{n-1}K_p \| + \| K^{n-1}K_p - K_p^n \| \\ &\leqq \| K \|^{n-1} \| K - K_p \| + \| K^{n-1} - K_p^{n-1} \| \cdot \| K_p \| \\ &\to 0 \qquad (p \to \infty), \end{aligned}$$

since $\| K_p \|$ is bounded. This disposes of (i).

The remaining statements of the lemma now follow rapidly. We have, first,

$$| \, \| K^n \| - \| K_p^n \| \, | \leqq \| K^n - K_p^n \| \to 0 \qquad (p \to \infty).$$

Next,

$$\begin{aligned} | \sigma_n - \sigma_{n,p} | &= | \tau(K^n) - \tau(K_p^n) | \\ &\to 0 \qquad (p \to \infty), \end{aligned}$$

by (i) and §2.7. Finally,

$$\begin{aligned} | (K^n x, y) - (K_p^n x, y) | &= | ((K^n - K_p^n)x, y) | \\ &\leqq \| K^n - K_p^n \| \cdot \| x \| \cdot \| y \| \\ &\to 0 \qquad (p \to \infty). \end{aligned}$$

5.4. Lemma. *Let K be of finite norm, $\sigma_n = \tau(K^n)$ $(n = 2, 3, \cdots)$. Define δ_n and Δ_n by the formulas of Theorems 4.4 and 4.5 respectively. Then*

$$\delta_0 = 1, \qquad \Delta_0 = I,$$

(5.4.1) $$\Delta_n = \delta_n I + K\Delta_{n-1} = \delta_n I + \Delta_{n-1}K \qquad (n = 1, 2, \cdots),$$

(5.4.2) $$\tau(\Delta_{n-1}K - \delta_{n-1}K) = \tau(K\Delta_{n-1} - \delta_{n-1}K) = -n\delta_n \qquad (n = 1, 2, \cdots).$$

Also

$$| \delta_n | \leqq \frac{e^{\frac{1}{2}n} \| K \|^n}{n^{\frac{1}{2}n}}, \qquad | \Delta_n | \leqq \frac{e^{\frac{1}{2}+\frac{1}{2}n} \| K \|^n}{n^{\frac{1}{2}n}}.$$

The identities are immediate. To prove the inequalities, we proceed as follows.

Define K_p $(p = 1, 2, \cdots)$ as in §5.2; if we regard K_p as a transformation in $\mathfrak{M}_p$, we can define $\delta_{n,p}$ and $\Delta_{n,p}$ $(n = 0, 1, 2, \cdots)$ in terms of the theory

of Chapter 4. We then extend the domain of definition of $\Delta_{n,p}$ to the whole of $\mathfrak{H}$ by the equation

$$\Delta_{n,p} = E_p \Delta_{n,p} E_p .$$

By Lemma 4.8,

$$|\delta_{n,p}| \leqq \frac{e^{\frac{1}{2}n} \| K_p \|^n}{n^{\frac{1}{2}n}} \leqq \frac{e^{\frac{1}{2}n} \| K \|^n}{n^{\frac{1}{2}n}} \qquad (n = 1, 2, \cdots);$$

and by Lemma 5.3, $\delta_{n,p} \to \delta_n$ $(p \to \infty)$. Hence

$$|\delta_n| \leqq \frac{e^{\frac{1}{2}n} \| K \|^n}{n^{\frac{1}{2}n}} \qquad (n = 1, 2, \cdots).$$

Thus the first inequality of the lemma is proved.

For arbitrary $x, y \in \mathfrak{H}$,

$$\begin{aligned} |(\Delta_{n,p} x, y)| &= |(E_p \Delta_{n,p} E_p x, y)| \\ &= |(\Delta_{n,p} E_p x, E_p y)| \\ &\leqq |\Delta_{n,p}| \cdot \| E_p x \| \cdot \| E_p y \|, \end{aligned}$$

where $|\Delta_{n,p}|$ is defined with respect to the domain $\mathfrak{M}_p$. Hence

$$|(\Delta_{n,p} x, y)| \leqq |\Delta_{n,p}| \cdot \| x \| \cdot \| y \|.$$

We also remark that, if

$$x = \sum_{\alpha=1}^{\infty} \xi_\alpha x_\alpha, \qquad y = \sum_{\alpha=1}^{\infty} \eta_\alpha x_\alpha,$$

then

$$(E_p x, y) = \sum_{\alpha=1}^{p} \xi_\alpha \bar{\eta}_\alpha \to \sum_{\alpha=1}^{\infty} \xi_\alpha \bar{\eta}_\alpha = (x, y) \qquad (p \to \infty).$$

It follows from this and from Lemma 5.3 that

$$(\Delta_{n,p} x, y) \to (\Delta_n x, y) \qquad (p \to \infty)$$

for all n. Now, by Lemma 4.9,

$$\begin{aligned} |(\Delta_{n,p} x, y)| &\leqq |\Delta_{n,p}| \cdot \| x \| \cdot \| y \| \\ &\leqq \frac{e^{\frac{1}{2}n+\frac{1}{2}} \| K_p \|^n}{n^{\frac{1}{2}n}} \| x \| \cdot \| y \| \\ &\leqq \frac{e^{\frac{1}{2}n+\frac{1}{2}} \| K \|^n}{n^{\frac{1}{2}n}} \| x \| \cdot \| y \|. \end{aligned}$$

Letting $p \to \infty$, we obtain

$$|(\Delta_n x, y)| \leqq \frac{e^{\frac{1}{2}n+\frac{1}{2}} \| K \|^n}{n^{\frac{1}{2}n}} \| x \| \cdot \| y \|$$

for arbitrary $x, y \in \mathfrak{H}$; i.e.,

$$|\Delta_n| \leqq \frac{e^{\frac{1}{2}n+\frac{1}{2}} \| K \|^n}{n^{\frac{1}{2}n}} \qquad (n = 1, 2, \cdots),$$

and this is the required result.

We now come to the main theorem of the paper.

5.5. THEOREM. *Let K be a transformation of finite norm, and let δ_n and Δ_n be defined as in Theorems 4.4 and 4.5. Put*

$$\delta(\lambda) = \sum_{n=0}^{\infty} \delta_n \lambda^n, \tag{5.5.1}$$

$$\Delta(\lambda) = \sum_{n=0}^{\infty} \Delta_n \lambda^n. \tag{5.5.2}$$

Then the series (5.5.1) is convergent for all λ, so that $\delta(\lambda)$ is an integral function of λ; for each value of λ, the series of transformations (5.5.2) is uniformly convergent, and $\Delta(\lambda)$ is a bounded transformation. Finally, if λ_0 is not a zero of $\delta(\lambda)$, the equation

$$(I - \lambda_0 K)x = y \tag{5.5.3}$$

has the unique solution

$$x = \frac{\Delta(\lambda_0)}{\delta(\lambda_0)} y. \tag{5.5.4}$$

The statements about the convergence of the series follow at once from the inequalities of Lemma 5.4. Also

$$\begin{aligned} |\Delta(\lambda)| &\leqq \sum_{n=0}^{\infty} |\Delta_n| \cdot |\lambda|^n \\ &\leqq \sum_{n=0}^{\infty} \frac{e^{\frac{1}{2}n+\frac{1}{2}} \| K \|^n |\lambda|^n}{n^{\frac{1}{2}n}}, \end{aligned}$$

so that $\Delta(\lambda)$ is a bounded transformation.

We now recall that Δ_n satisfies the recurrence relations

$$\Delta_0 = I, \qquad \Delta_n = \delta_n I + K\Delta_{n-1} = \Delta_n I + \Delta_{n-1} K \qquad (n = 1, 2, \cdots).$$

Hence

$$\begin{aligned} (I - \lambda_0 K)\Delta(\lambda_0) &= \sum_{n=0}^{\infty} (I - \lambda_0 K)\Delta_n \lambda_0^n \\ &= \sum_{n=0}^{\infty} (\Delta_n - \lambda_0 \Delta_{n+1} + \lambda_0 \delta_{n+1} I)\lambda_0^n \\ &= I\,(1 + \sum_{n=1}^{\infty} \delta_n \lambda_0^n) \\ &= \delta(\lambda_0) I. \end{aligned}$$

Similarly $\Delta(\lambda_0)(I - \lambda_0 K) = \delta(\lambda_0) I$.

Now, if $\delta(\lambda_0) \neq 0$, and

$$x = \frac{\Delta(\lambda_0)}{\delta(\lambda_0)} y,$$

we have $(I - \lambda_0 K)x = (I - \lambda_0 K)\Delta(\lambda_0)y/\delta(\lambda_0) = y$; i.e., x is a solution of (5.5.3). On the other hand, if $\delta(\lambda_0) \neq 0$, and

$$y = (I - \lambda_0 K)x,$$

then

$$\frac{\Delta(\lambda_0)}{\delta(\lambda_0)} y = \Delta(\lambda_0)(I - \lambda_0 K)x/\delta(\lambda_0) = x;$$

i.e., $[\delta(\lambda_0)]^{-1}\Delta(\lambda_0)y$ is the unique solution of (5.5.3). This completes the proof of the theorem.

It should be noted that the recurrence formula (5.4.1), which ensures that (5.5.4) gives a solution of (5.5.3), is proved without using the fact that $\sigma_n = \tau(K^n)$. It follows from this that the actual values of the constants σ_n are immaterial, provided only that they are chosen in such a way that the series (5.5.1) and (5.5.2) are convergent for all values of λ; what we have shown is that this can be brought about by taking $\sigma_n = \tau(K^n)$ $(n = 2, 3, \cdots)$. Nevertheless, the field of choice for the sequence $\{\sigma_n\}$ is not so wide as one might imagine; for, if the series (5.5.2) is to converge for all values of λ, the zeros of $\delta(\lambda)$ must cancel all the poles of $\Delta(\lambda)/\delta(\lambda)$, and the latter expression is independent of the way in which we choose the sequence $\{\sigma_n\}$. It is, however, possible, for instance, to replace any finite set of these constants by zeros, or by other constants chosen at random, without affecting the convergence of either series.

We shall now discuss the convergence of the series (5.5.2) rather more closely.

5.6. THEOREM. *For all values of* λ,

$$\Delta(\lambda) = \delta(\lambda)I + \lambda H(\lambda),$$

where

$$H(\lambda) = \sum_{n=0}^{\infty} H_n \lambda^n, \tag{5.6.1}$$

$$H_n = K\Delta_n = \Delta_n K \qquad (n = 0, 1, 2, \cdots),$$

each H_n *is of finite norm,* $H(\lambda)$ *is of finite norm, and the series* (5.6.1) *is convergent in norm.*

We have

$$\Delta_n = \delta_n I + K\Delta_{n-1} \qquad (n = 1, 2, \cdots).$$

Hence

$$\begin{aligned}\Delta(\lambda) &= \sum_{n=0}^{\infty} \Delta_n \lambda^n \\ &= \Big(\sum_{n=0}^{\infty} \delta_n \lambda^n\Big) I + \sum_{n=1}^{\infty} K\Delta_{n-1}\lambda^n \\ &= \delta(\lambda) I + \lambda \sum_{n=0}^{\infty} H_n \lambda^n,\end{aligned}$$

the series being uniformly convergent for each λ.

Now, by Lemma 2.6,

$$\begin{aligned}\| \lambda^n H_n \| &= \| \lambda^n K \Delta_n \| \\ &\leqq | \lambda |^n | \Delta_n | \cdot \| K \| \\ &\leqq \frac{e^{\frac{1}{2}n+\frac{1}{2}} \| K \|^{n+1} | \lambda |^n}{n^{\frac{1}{2}n}};\end{aligned}$$

the series $\sum_{n=0}^{\infty} \| \lambda^n H_n \|$ is therefore convergent, so that $\sum_{n=0}^{\infty} H_n \lambda^n$ is convergent in norm, and $H(\lambda)$ is of finite norm. This completes the proof.

Since H_n is of finite norm, the equation $y = H_n x$ can be written

$$y(s) = \int h_n(s, t) x(t)\, dt,$$

where $\iint | h_n(s, t) |^2 \, ds\, dt < \infty$. We shall now use this fact to tighten up our results on the convergence of the series for the solution still further.

5.7. Theorem. *If $\delta(\lambda_0) \neq 0$, the solution of the equation*

$$x(s) = y(s) + \lambda_0 \int k(s, t) x(t)\, dt \tag{5.7.1}$$

can be written in the form

$$x(s) = y(s) + \frac{\lambda_0}{\delta(\lambda_0)} \int h(s, t; \lambda_0) y(t)\, dt, \tag{5.7.2}$$

where

$$h(s, t; \lambda) = \sum_{n=0}^{\infty} \lambda^n h_n(s, t), \tag{5.7.3}$$

the series being convergent for almost all (s, t); furthermore,

$$\iint | h_n(s, t) |^2 ds\, dt < \infty \qquad (n = 0, 1, 2, \cdots),$$

and, for almost all (s, t),

$$h_0(s, t) = k(s, t),$$

$$| h_1(s, t) | \leqq e^{\frac{1}{2}} \| K \| \cdot | k(s, t) | + a(s)b(t),$$

$$| h_n(s, t) | \leqq \frac{e^{\frac{1}{2}n} \| K \|^n}{n^{\frac{1}{2}n}} | k(s, t) | + \frac{e^{\frac{1}{2}n} \| K \|^{n-1}}{(n-1)^{\frac{1}{2}(n-1)}} a(s)b(t) \qquad (n = 2, 3, \cdots), \tag{5.7.4}$$

where

$$a(s) = \left\{ \int | k(s, t) |^2 dt \right\}^{\frac{1}{2}}, \qquad b(t) = \left\{ \int | k(s, t) |^2 ds \right\}^{\frac{1}{2}}.$$

By Theorem 5.6,

$$H(\lambda) = \sum_{n=0}^{\infty} H_n \lambda^n,$$

where

$$H_0 = K, \qquad H_n = K\Delta_n = \delta_n K + K\Delta_{n-1}K \qquad (n = 1, 2, \cdots).$$

Hence

$$h_n(s, t) = \delta_n k(s, t) + g_n(s, t),$$

where $g_0(s, t) = 0$, and, when $n \geqq 1$, $g_n(s, t)$ is the kernel of the transformation $K\Delta_{n-1}K$.

By Lemma 2.9,

$$| g_n(s, t) | \leqq | \Delta_{n-1} | \cdot a(s)b(t) \qquad (n = 1, 2, \cdots)$$

for almost all (s, t). The inequalities (5.7.4) now follow from Lemma 5.4. The series (5.7.3) is therefore convergent for almost all (s, t); since its partial sums are dominated by an expression of the form

$$A[| k(s, t) | + a(s)b(t)],$$

the fact that (5.7.2) gives a solution of (5.7.1) can indeed be verified directly by term-by-term integration.

We conclude by showing that the formulas obtained in this paper are equivalent to the modified Fredholm formulas introduced by Hilbert to deal with certain discontinuous kernels and used by Carleman in the general case. (Cf. Hilbert [6] and Carleman [1].)

5.8. Theorem. *Let* $h_n(s, t)$ *be the kernel defined in Theorem* 5.7. *Then*

$$h_n(s, t) = \frac{(-1)^n}{n!} \int \cdots \int N_n^*(s, t)\, du_1 \cdots du_n, \tag{5.8.1}$$

and

$$\delta_n = \frac{(-1)^n}{n!} \int \cdots \int N_n du_1 \cdots du_n \qquad (n = 1, 2, \cdots), \tag{5.8.2}$$

the N's being defined as in §1.1.

Denote the expressions on the right sides of (5.8.1) and (5.8.2) by $g_n(s, t)$ and ϵ_n respectively, and let G_n be the transformation whose kernel is $g_n(s, t)$. Clearly

$$g_1(s, t) = \int k(s, u)k(u, t)\,du = h_1(s, t),$$

$$\epsilon_1 = 0 = \delta_1.$$

Now we have, for $n = 2, 3, \cdots$,

$$H_n = K\Delta_n = K(\delta_n I + K\Delta_{n-1}) = \delta_n K + KH_{n-1},$$

$$\delta_n = -\frac{1}{n}\tau(K\Delta_{n-1} - \delta_{n-1}K) = -\frac{1}{n}\tau(H_{n-1} - \delta_{n-1}K).$$

We shall show that G_n and ϵ_n satisfy the same recurrence formulas.

The formula

$$\epsilon_n = -\frac{1}{n}\tau(G_{n-1} - \epsilon_{n-1}K)$$

is obvious by inspection.

Secondly, we have

$$g_n(s, t) = \epsilon_n k(s, t) + \frac{(-1)^n}{n!}\int \cdots \int N_n^*(s, t)\,du_1 \cdots du_n$$

$$= \epsilon_n k(s, t) + \frac{(-1)^n}{n!}\,q_n(s, t),$$

say. To evaluate $q_n(s, t)$, we expand the determinant by its top row. As readily seen, all terms in the expansion of $q_n(s, t)$ are equal so that

$$g_n(s, t) = \epsilon_n k(s, t) + \int k(s, u)g_{n-1}(u, t);$$

that is,

$$G_n = \epsilon_n K + KG_{n-1}.$$

This is the required result.

Consequently we have $\epsilon_1 = \delta_1$, $G_1 = H_1$, and G_n, ϵ_n satisfy the same recurrence formulas as H_n, δ_n. We must therefore have $H_n = G_n$, $\delta_n = \epsilon_n$ for all n, and the theorem is proved.

References

1. T. Carleman, *Zur Theorie der linearen Integralgleichungen*, Mathematische Zeitschrift, vol. 9(1921), pp. 196–217.
2. R. Courant and D. Hilbert, *Methoden der mathematischen Physik*, vol. 1, 2d ed., Berlin, 1931.
3. I. Fredholm, *Sur une classe d'équations fonctionnelles*, Acta Mathematica, vol. 27(1903), pp. 365–390.
4. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge, 1934.
5. E. Hellinger and O. Toeplitz, *Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten*, Encyklopädie der mathematischen Wissenschaften, vol. 2(1927), part 3, 2d half, pp. 1335–1601.
6. D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen* (Part 1), Nachrichten der Gesellschaft der Wissenschaften zu Göttingen, 1904, pp. 49–91.
7. E. Hille and J. D. Tamarkin, *On the characteristic values of linear integral equations*, Proceedings of the National Academy of Sciences, vol. 14(1928), pp. 911–914.
8. E. Hille and J. D. Tamarkin, *On the characteristic values of linear integral equations*, Acta Mathematica, vol. 57(1931), pp. 1–76.
9. G. Julia, *Introduction Mathématique aux Théories Quantiques*, part 2, Paris, 1938.
10. A. D. Michal and R. S. Martin, *Some expansions in vector space*, Journal de Mathématiques pures et appliquées, (9), vol. 13(1934), pp. 69–91.
11. M. H. Stone, *Linear Transformations in Hilbert Space and Their Applications to Analysis*, New York, 1932.
12. E. C. Titchmarsh, *The Theory of Functions*, Oxford, 1932.

St. John's College, Cambridge, England.

PSEUDO-NORMED LINEAR SPACES

By J. P. LaSalle

Hyers [2][1] introduced the concept of pseudo-normed linear spaces (p.l.s.'s) and showed that such spaces are equivalent to linear topological spaces (l.t.s.'s). Throughout the present paper we shall deal with p.l.s.'s, as it is in general more convenient, though it is to be remembered that what we prove for p.l.s.'s applies also to l.t.s.'s. The terms used which are not defined are those for l.t.s.'s (see [3]). For instance, when we speak of a convex p.l.s., we mean that the l.t.s. that is equivalent to the p.l.s. is convex. In this paper a necessary and sufficient condition is given that there exist a non-null linear functional on a p.l.s., and also it is shown that the set of all linear functions on a p.l.s. to any other p.l.s. is itself a p.l.s. It is then shown by example[2] that there do exist p.l.s.'s on which no non-null linear functional can be defined.

1. Let T with elements $x, y, \cdots$ be a p.l.s. with respect to a strongly partially ordered space [2] D; that is, T is a linear space such that there exists a real-valued function $n(x, d)$ defined on TD which satisfies the following postulates:

(1) $n(x, d) \geqq 0$; $n(x, d) = 0$ for all $d \in D$ implies $x = \theta$, where θ is the zero element of T;

(2) $n(\alpha x, d) = |\alpha| \, n(x, d)$ for all $x \in T$, $d \in D$, and α real;

(3) given $\eta > 0$, $e \in D$, there exist $\delta > 0$, $d \in D$ such that $n(x + y, e) < \eta$ for $n(x, d) < \delta$ and $n(y, d) < \delta$;

(4) $d > e$ implies that $n(x, d) \geqq n(x, e)$.

$n(x, d)$ is called the pseudo-norm of x with respect to d.

Let αd represent the association of a positive real number α with an element $d \in D$. Define[3] $1 \cdot e = e$, $\alpha(\beta e) = (\alpha\beta)e = (\beta\alpha)e$, $E = [\alpha d;\, d \in D,\, \alpha > 0]$ and $n(x, \alpha d) = \alpha n(x, d)$, for $\alpha, \beta > 0$. Then E is a strongly partially ordered space with $e_1 \geqq e_2$, $e_1, e_2 \in E$, if $n(x, e_1) \geqq n(x, e_2)$ for all $x \in T$, and $e_1 = e_2$ if $n(x, e_1) = n(x, e_2)$ for all $x \in T$. This is consistent with the definitions already given. $n(x, e)$ is a pseudo-norm of x with respect to e; that is, $n(x, e)$ satisfies postulates (1)–(4). This modified pseudo-norm gives a more convenient statement of (3), namely,

(3') given $e \in E$ there exist $f \in E$ such that $n(x + y, e) \leqq n(x, f) + n(y, f)$ for all $x, y \in T$;

Received August 5, 1940. The author wishes to thank Professor Michal for his help and guidance.

[1] The numbers in brackets refer to the bibliography at the end of the paper.

[2] The author is indebted to the referee for this example. For another example see Theorem 1 of [1]. A proof of Theorem 1 of [1] can be given if we use Theorem 4 of this paper.

[3] $[x;\ ]$ denotes the set of all x's having the property following the semicolon.

and in the remainder of the paper T shall denote a p.l.s. as modified above. Also for $x \in T$, $e \in E$, $U_x(e)$ is defined by $U_x(e) = [y; n(x - y, e) < 1]$, and $U_\theta(e)$ is denoted by just $U(e)$.[4]

We can immediately state the following theorem:

THEOREM 1. *A necessary and sufficient condition that a p.l.s. T be locally bounded is that there exist an $e_0 \in E$ such that given $e \in E$ there exists a real number $\alpha > 0$ such that $\alpha e_0 \geqq e$.*

By virtue of Theorem 1 and a theorem by Hyers,[5] Kolmogoroff's normability criterion may be stated in the obvious manner.

THEOREM 2. *A necessary and sufficient condition that a p.l.s. T be normable is that there exist an $e_0 \in E$ satisfying the condition of Theorem* 1 *and in addition $n(x + y, e_0) \leqq n(x, e_0) + n(y, e_0)$ for all $x, y \in T$.*

2. THEOREM 3. *A necessary and sufficient condition that there exist a non-null linear functional defined on a p.l.s. T is that there exist a convex open set containing the zero element and properly contained in T.*

Proof. Necessity. Let $F(x)$ be a non-null linear functional which exists by hypothesis. Then by Theorem 1.3 of [3] there exists an $e \in E$ such that $|F(x)| \leqq n(x, e)$. Define $S = [x; |F(x)| < 1]$. Then given $x_0 \in S$, define $\delta = 1 - |F(x_0)|$, and $\delta > 0$, since $F(x_0) < 1$. $|F(x)| \leqq |F(x_0 - x)| + |F(x_0)|$, since $F(x)$ is additive, and $|F(x)| \leqq n(x_0 - x, e) + |F(x_0)|$. Consider $x \in U_{x_0}(\delta^{-1}e)$, that is, $n(x_0 - x, e) < \delta$. Then $|F(x)| < \delta + |F(x_0)| = 1$ and $x \in S$, that is, $U_{x_0}(\delta^{-1}e) \subset S$; and hence by Theorem 1.2 of [3], S is open. θ is evidently contained in S, and since $F(x)$ is non-null, there exist $x \in T$, $x \bar{\in} S$.

Consider $x \in (1 - \alpha)S + \alpha S$ $(0 < \alpha < 1)$; that is, $x = (1 - \alpha)x_1 + \alpha x_2$, $x_1, x_2 \in S$. Then $|F(x)| \leqq (1 - \alpha)|F(x_1)| + \alpha |F(x_2)| < 1$, and hence $(1 - \alpha)S + \alpha S \subset S$, and S is convex. Therefore S is an open convex set containing the zero element and properly contained in T, and the necessity is proved.

Sufficiency. Let S be the convex open set that we assume to exist, $\theta \in S$, $S \neq T$. Since θ is contained in S, we see by Theorem 1.2 of [3] that there exists an $e \in E$ such that $U(e) \subset S$. Hence given $x \in T$ we see by (2) of the definition of the pseudo-norm that, for $\lambda > n(x, e)$, $n(x/\lambda, e) < 1$. That is, $x/\lambda \in U(e)$ or $x \in \lambda U(e) \subset \lambda S$.

Define $F(x) = \text{g.l.b.} [\lambda; \lambda > 0, x \in \lambda S]$. We see that $F(x)$ is then defined for each $x \in T$. Moreover, $F(x)$ has the following properties:

(a) $F(\alpha x) = \alpha F(x)$, $\alpha \geqq 0$, $x \in T$.

(b) $F(x + y) \leqq F(x) + F(y)$ for all $x, y \in T$.

(c) $F(x) \leqq n(x, e)$ for all $x \in T$.

[4] The set of all $U(e)$, $e \in E$, is equivalent to a fundamental neighborhood system of T as a l.t.s. See [2], Theorem 1.

[5] See [2], Theorem 2. If $n(x, e)$ satisfies the triangular inequality, then $U(e)$ is convex and also open.

In order to show that $F(x)$ satisfies (a) let $x \in \lambda S$, $\lambda > 0$. Then $\alpha x \in \alpha\lambda S$. Hence, if $\alpha > 0$, then $F(\alpha x) \leqq \alpha F(x)$. Conversely, if $\alpha x \in \lambda S$ for $\lambda > 0$, $\alpha > 0$, then $x \in \lambda\alpha^{-1}S$, and $F(x) \leqq \alpha^{-1}F(\alpha x)$. Then we have also $\alpha F(x) \leqq F(\alpha x)$, and hence $F(\alpha x) = \alpha F(x)$ for $x \in T$, $\alpha > 0$. If $\alpha = 0$, then $F(\alpha x) = \theta$, and (a) is satisfied by $F(x)$.

Next consider (b). If $x \in \alpha S$, $y \in \beta S$, $\alpha > 0$, $\beta > 0$; then, since S is convex, $(\alpha + \beta)^{-1}(x + y) \in \alpha(\alpha + \beta)^{-1}S + \beta(\alpha + \beta)^{-1}S \subset S$, and $x + y \in (\alpha + \beta)S$. Hence $F(x + y) \leqq F(x) + F(y)$.

Finally, we show that $F(x)$ has property (c). As we have shown before, for every $\lambda > n(x, e)$, $x \in \lambda U(e) \subset \lambda S$, and hence $F(x) \leqq \lambda$ for every $\lambda > n(x, e)$. Therefore $F(x) \leqq n(x, e) + \epsilon$ for every $\epsilon > 0$, or $F(x) \leqq n(x, e)$ for all $x \in T$.

There exist by hypothesis $x \in T$, $x \bar{\in} S$, and for such x, $F(x) > 0$. The sufficiency then follows from Theorem 1.9 of [3], which is a generalization of a well-known theorem of Banach.

LEMMA 1. *Given $e \in E$, there exists an open set which contains the zero element and is contained in $U(e)$.*

Proof. By (3′) there exists an $f \in E$ such that $n(x + y, e) \leqq n(x, f) + n(y, f)$. Let $x \in U(f)$, so that $n(x, f) < 1$, and let $\delta = 1 - n(x, f)$. Then, if $y \in U_x(\delta^{-1}f)$, $n(y, e) \leqq n(x, f) + n(x - y, f) < n(x, f) + \delta = 1$; that is, $U_x(\delta^{-1}f) \subset U(e)$, and x is not a limit element (see [3]) of the complement of $U(e)$. Hence since $I(U(e)) = U(e) \cdot C((CU(e))')$ (see [4], p. 4), $U(f) \subset I(U(e))$,[6] and the lemma is proved.

LEMMA 2. *The convex hull of an open set is an open set.*

Proof. Let S be the open set and denote the convex hull of S by S_{conv}. Let $x_0 \in S_{conv}$; that is, $x_0 = \sum_{i=1}^{n} \alpha_i x_i$, $\alpha_i > 0$, $x_i \in S$, $\sum_{i=1}^{n} \alpha_i = 1$. Since S is open, we see by Theorem 1.2 of [3] that for each x_i there exists a $U_{x_i}(e_i) \in S$. Hence $\sum_{i=1}^{n} \alpha_i U_{x_i}(e_i) \subset S_{conv}$, and since $U_x(e) = x + U(e)$ for all $x \in T$, $e \in E$, we have $\sum_{i=1}^{n} \alpha_i U_{x_i}(e_i) = x_0 + \sum_{i=1}^{n} \alpha_i U(e_i)$. Therefore $U_{x_0}(\alpha_i^{-1}e_i) \subset S_{conv}$ $(i = 1, 2, \cdots, n)$, and by Theorem 1.2 of [3] S_{conv} is open.

By the above two lemmas and Theorem 3, we may prove the following theorem.

THEOREM 4. *A necessary and sufficient condition that there exist no non-null linear functional on a p.l.s. T is that, for every $e \in E$, $U(e)_{conv} = T$; i.e., given $e \in E$, every element $x \in T$ can be represented as $x = \sum_{i=1}^{n} \alpha_i x_i$, where $\alpha_i > 0$, $\sum_{i=1}^{n} \alpha_i \leqq 1$, and $n(x_i, e) < 1$ for $i = 1, 2, \cdots, n$.*

[6] CS denotes the complement of S, $I(S)$ denotes the interior of S, and S' denotes the derived set of S.

Proof. Necessity. Let no non-null linear functional exist on T. Then assume that there exist an $e \in E$ and an $x \in T$ such that $x \bar{\in} U(e)_{\text{conv}}$. Now by Lemma 1 there exists an open set $S_\theta \subset U(e)$, $\theta \in S_\theta$. Then $(S_\theta)_{\text{conv}}$ is by Lemma 2 a convex open set and $x \bar{\in} (S_\theta)_{\text{conv}} \subset U(e)_{\text{conv}}$. But then by Theorem 3 there does exist a non-null linear functional on T. This is a contradiction, and the necessity is proved.

Sufficiency. Let $U(e)_{\text{conv}} = T$ for every $e \in E$. Then assume that there exists a non-null linear functional defined on T. But then by Theorem 3 there exists a convex open set $S_\theta \subset T$ such that $\theta \in S_\theta$, $S_\theta \neq T$. Since $\theta \in S_\theta$, we see by Theorem 1.2 of [3] that there exists an $e \in E$ such that $U(e) \subset S_\theta$. Hence $U(e)_{\text{conv}} \subset S_\theta \neq T$, since S_θ is convex. This is a contradiction, and the sufficiency is proved.

An example (see footnote 2) of a p.l.s. on which there exists no non-null linear functional is the space of measurable functions on the real open interval (0, 1) as given by Hyers ([2], p. 624). Let D be the open interval (0, 1) with ordering as usual, and the pseudo-norm $n(x, d)$ is defined by

$$n(x, d) = \underset{m(S) \geqq d}{\text{g.l.b.}} \; [\underset{t \in S}{\text{l.u.b.}} \; | x(t) |]$$

for $x \in T$, $d \in D$, and S a measurable subset of (0, 1).

Let x be any element of T and d any element of D, that is, $0 < d < 1$. Let n be the smallest integer such that $n > (1 - d)^{-1}$ $(n \geqq 2)$. Define $x_i(t) = nx(t)$ for $(i - 1)/n \leqq t \leqq i/n$ and 0 otherwise. Since $x_i(t)$ is zero on a set of measure $1 - n^{-1} > d$, $n(x_i, d) = 0 < 1$, for $i = 1, 2, \cdots$, while $\sum_{i=1}^{n} n^{-1} x_i(t) = x(t)$, $\sum_{i=1}^{n} n^{-1} = 1$. Hence $x(t) \in U(d)_{\text{conv}}$, and therefore $U(d)_{\text{conv}} = T$ for all $d \in D$. Evidently this holds if we associate positive real numbers with the elements of D, since $n(x_i, \alpha(d)) = \alpha n(x_i, d) = 0$, for $\alpha > 0$. Therefore, by Theorem 4, T is a p.l.s. such that there exists no non-null linear functional on T.

4. A set $S \subset T$ is said to be *bounded* if given $e \in E$ there exists a number $\alpha(e)$ such that $n(x, e) \leqq \alpha(e)$ for all $x \in S$. Let B_T be the set of all bounded sets S in T such that $\theta \in S$. If $S_1 \in B_T$, $S_2 \in B_T$, then $S_1 + S_2 \in B_T$, $\alpha S_1 \in B_T$, α real, and every finite set containing the origin is in B_T. B_T is clearly a strongly partially ordered space with $S_1 \geqq S_2$ if $S_2 \subset S_1$.

Let T and T' be p.l.s.'s with respect to E and E', and let F be the linear set of all linear functions on T to T', let ϑ represent the zero element of F, i.e., the null function on T to T'. For $\Phi \in F$ it evidently follows by Theorem 1.3 of [3] that $\Phi(x)$ carries bounded sets into bounded sets. Define W to be the set of all pairs $w = (e', S)$, where $e' \in E'$ and $S \in B_T$. Since E' and B_T are strongly partially ordered spaces, so is W with $w_1 > w_2$ if $e_1' > e_2'$, $S_1 \geqq S_2$. Define $M(\Phi, w) = \underset{x \in S}{\text{l.u.b.}} \; n(\Phi(x), e')$, which exists for every $\Phi \in F$ and every $w \in W$.

THEOREM 5. *F is a p.l.s. with respect to the strongly partially ordered space W and with pseudo-norm $M(\Phi, w)$.*

Proof. In order to prove the theorem we must show that $M(\Phi, w)$ satisfies (1)–(4) of §1.

$M(\Phi, w)$ is defined for all $\Phi \epsilon F$ and $w \epsilon W$; $M(\Phi, w) \geqq 0$ by definition; and by (1) of the definition of a p.l.s. $M(\Phi, w) = 0$ for all $w \epsilon W$ implies $\Phi = \vartheta$, since given $x \epsilon T$ there exists an $S \epsilon B_T$ such that $x \epsilon S$. Thus $M(\Phi, w)$ satisfies (1) of §1.

Next consider (2). Clearly $M(\alpha\Phi, w) = |\alpha| \cdot M(\Phi, w)$. Also we can define $\alpha w = (\alpha e', S) = (e', \alpha S)$ for $\alpha > 0$ and obtain $M(\Phi, \alpha w) = \alpha M(\Phi, w)$.

(3′) follows from (3′) for the property of $n(\Phi(x), e')$.

(4) is evident from the definition of $M(\Phi, w)$ and property (4) of the pseudo-norm $n(\Phi(x), e')$.

THEOREM 6. *If T' is convex, i.e., the pseudo-norm satisfies the triangular inequality for all $e' \epsilon E'$* (see footnote 5), *then F is convex.*

Proof. This is clear since then $n(\Phi_1(x) + \Phi_2(x), e') \leqq n(\Phi_1(x), e') + n(\Phi_2(x), e')$, and hence by definition $M(\Phi_1 + \Phi_2, w) \leqq M(\Phi_1, w) + M(\Phi_2, w)$.

COROLLARY 1. *The set of all linear functionals on T, a p.l.s., is a convex p.l.s. with respect to the strongly partially ordered set B_T.*

COROLLARY 2. *If there exists a bounded open set S in T, $\theta \epsilon S$, then the set of all linear functionals on T is a normed linear space.*

Proof. This follows from Corollary 1 and Theorem 2, since B_T then satisfies the conditions of Theorem 2.

BIBLIOGRAPHY

1. M. M. DAY, *The spaces L^p with $0 < p < 1$*, Bulletin of the American Mathematical Society, vol. 46(1940), pp. 816–823.

2. D. H. HYERS, *Pseudo-normed linear spaces and Abelian groups*, this Journal, vol. 5(1939), pp. 628–634.

3. J. P. LASALLE, *Application of the pseudo-norm to the study of linear topological spaces*, to be published in the July issue of the Revista des Ciencias, University of Lima.

4. W. SIERPINSKI, *Introduction to General Topology* (translation), Toronto, 1934.

CALIFORNIA INSTITUTE OF TECHNOLOGY.

MONOTONE TRANSFORMATIONS OF NON-COMPACT TWO-DIMENSIONAL MANIFOLDS

By Venable Martin

1. **Introduction.** The object of this paper is to give an extension of results of Roberts and Steenrod.[1] The generalization in Part I over the results of Roberts and Steenrod consists in removing the assumption of compactness on the manifold, but replacing it by the assumption that the continua are compact. In this case we obtain a complete characterization of the image space of the manifold. Specifically, we define an A-space with identifications (see §§7 and 10 below) and prove that if G is an upper semi-continuous collection of compact continua filling a 2-manifold, then G is an A-space with identifications; and conversely, if S is an A-space with identifications, then there is a 2-manifold M and an upper semi-continuous collection G of compact continua filling M such that G is homeomorphic to S (see Theorem 5, Part I). In addition, we are able to prove under more restrictive hypotheses that G is a manifold (see Theorems 1 and 2), and under still different hypotheses that G is an A-space without identifications (see Theorems 3 and 4).

In Part II neither the manifold nor the continua are assumed to be compact, but the image space is assumed to be metric and the characterization of the image space is effected only in the case in which the manifold has a finite 1-dimensional Betti number. Moreover, we do not show that the characterization is complete in the sense that any space satisfying the restrictions of the characterization is a monotone image of a 2-manifold.

Part I

2. **Notation.** Throughout this paper M will denote a 2-dimensional manifold without boundary; G will denote an upper semi-continuous collection of continua filling M; G will also be used to denote the topological space whose points are the continua of this collection, with an element g of G defined as a limit element of a sequence of elements $g_1, g_2, \cdots$ of G if and only if there is a point of g which is a limit point of the point set $g_1 + g_2 + \cdots$, where $g_i \neq g$

Received September 2, 1940; presented in part to the American Mathematical Society, December 29, 1938. The author wishes to acknowledge the assistance of Professor J. H. Roberts in the writing of this paper.

[1] J. H. Roberts and N. E. Steenrod, *Monotone transformations of two-dimensional manifolds*, Annals of Mathematics, vol. 39(1938), pp. 851–862. This paper will be cited as MT. As part of the introduction to the present article we assume a reading of the introduction and of the statements of the lemmas and theorems of MT. In particular in Part I we shall speak of upper semi-continuous collections of continua or of monotone transformations as is convenient, without always calling attention to the equivalence of the two points of view.

$(i = 1, 2, \cdots)$; g or g_n will denote an element of G; g_p will denote that element of G which contains the point p of M; $R(S)$ will denote the mod 2 1-dimensional Betti number of the set S.

When we choose to view the problem as a problem of monotone transformations, we shall let G denote a metric space onto which M is mapped by a monotone transformation f. In this case the upper semi-continuous collection filling M, each element of which is an inverse image set under f of a point in G, will be referred to as *the collection to which f gives rise* and will again be denoted by G. This need cause no confusion when it is remembered that the upper semi-continuous collection is homeomorphic to the image space.[2] In Part I the continua of G will be assumed to be compact; in Part II some or all of the continua of G may be non-compact except where the context clearly implies the contrary. By a 1-sphere we mean a simple closed curve; M_n^0 will denote the interior of the manifold-with-boundary M_n. We shall use σ to denote the metric function in the space M, and ρ to denote the metric in G; $S_G(g, a)$ and $S_M(p, a)$ will denote, respectively, the sphere in G of radius a about g and the sphere in M of radius a about p; if no ambiguity is possible, the subscripts M and G may be omitted.

3. Lemma 1. *If M is non-compact, then $M = M_1 + M_2 + \cdots$, where for every n M_n is a compact manifold-with-boundary, the boundary being a finite set of mutually exclusive* 1-*spheres, $C_1^n, C_2^n, \cdots, C_{k_n}^n$, and the following properties hold*:

(1) $C_i^n C_j^k = 0$ *if either $i \neq j$ or $n \neq k$*;

(2) M_{n+1}^0 *contains* M_n;

(3) *for every $n > 1$, no proper subset of the set of* 1-*spheres $C_1^n, C_2^n, \cdots, C_{k_n}^n$ bounds a compact manifold-with-boundary which contains M_{n-1} in its interior*;

(4) *for every i and every n, C_i^n separates M; and*

(5) *if Y is a compact subset of M, then there exists an integer n such that M_n contains Y.*

The proof of this lemma is omitted.[3]

We shall sometimes have occasion to speak of a decomposition of M into $M_1 + M_2 + \cdots$ satisfying the first three and the fifth properties of Lemma 1 but not necessarily satisfying the fourth property. This weaker form of the lemma we shall call Lemma 1′.

We now state without proof the following lemma which is intuitively acceptable and which, in view of Lemmas 1 and 1′, follows without great difficulty from a Principal Theorem on the topology of manifolds.[4]

[2] See footnote 1.

[3] See B. v. Kerékjártó, *Vorlesungen über Topologie*, Berlin, 1923. The proof follows immediately from § 5, pp. 172, 173 of this book. Condition (5) of the lemma is a consequence of conditions (1), (2) and (3), but we find it convenient to formulate the lemma as above.

[4] See Kerékjártó, loc. cit., pp. 170, 171.

LEMMA A. *Suppose that*

(1) *G is an upper semi-continuous collection of continua filling the manifold M, and G is itself a manifold;*

(2) *$M = W_1 + W_2 + \cdots$ and $G = G_0 + G_1 + G_2 + \cdots$, where these decompositions are in accordance with Lemma* 1′;

(3) *$M = W_3 + W_6 + \cdots + W_{3n} + \cdots$ is a decomposition of M in accordance with Lemma* 1;

(4) *for $n = 0, 1, 2, \cdots$, G_n is homeomorphic both to W_{3n+1} and to W_{3n+2};*

(5) *W_{3n+2} contains G_n and G_n contains W_{3n+1};*

(6) *$W^0_{3n+2} - W_{3n+1}$ consists of a finite number of open cylinders; and*

(7) *if T_n is a component of $W_{3n+5} - W_{3n+1}$, then that component of $G_{n+1} - G_n$ which is contained in T_n is homeomorphic to T_n.*

Then it follows that M and G are homeomorphic.

LEMMA 2. *Given any integer i, there exists an integer n_i such that if $g \cdot M_i \neq 0$, then $M^0_{n_i}$ contains g.*

Proof. Suppose the lemma is false. Then there exists an integer i such that for every integer n there is an element g_n such that M^0_n does not contain g_n and $g_n \cdot M_i \neq 0$. Then there is a point p in M_i which is a limit point of $g_1 + g_2 + \cdots$, but which is not a point of $g_1 + g_2 + \cdots$, and hence g_p, the element which contains p, is a limit element of $g_1 + g_2 + \cdots$. Choose k so that M^0_k contains g_p. Then by the definition of upper semi-continuity, g_p is a limit element of those elements of the sequence $g_1, g_2, \cdots$ each of which lies in M^0_k. But there is only a finite number of such elements, and this contradiction proves the lemma.

4. **Notation.** Throughout the rest of Part I, n_i will denote the smallest integer having the property stated in Lemma 2. Likewise, if A is a compact set, n_A will denote the smallest integer such that if $g \cdot A \neq 0$, then $M^0_{n_A}$ contains g; and if g is an element, n_g will denote the smallest integer such that $M^0_{n_g}$ contains g. Further, M'_i will denote the compact manifold without boundary obtained from M_i by fitting 2-cells onto the boundary curves of M_i. Finally, by an *integral* subset of M we shall mean a subset A of M which has the property that if an element of G intersects A, then it lies wholly in A.

5. THEOREM 1. *If G contains at least two elements and $R(g) = 0$ for each g of G, then G is homeomorphic to M.*

Proof. The method of proof will be to construct decompositions of M and G, $M = W_1 + W_2 + \cdots$, $G = G_0 + G_1 + \cdots$, which satisfy the conditions of Lemma A. It will then follow by Lemma A that M and G are homeomorphic. As the first step in the proof we note that G is a manifold. To prove this we need only show that G is locally a plane at each of its elements. But this is proved exactly as in the proof of Theorem 1 of MT.

Now let $M = M_1 + M_2 + \cdots$ in accordance with Lemma 1. Let A_1 denote

the set of those elements of G each of which intersects a boundary curve of M_1. A_1 is a closed and compact set. Let B_1 denote A_1 plus those elements of G each of which lies in a compact component of $M - A_1$. Then B_1 is a compact continuum containing M_1. Now $M - B_1$ has only a finite number of components, $D_1, D_2, \cdots, D_p$. Modifying slightly the notation of Lemma 1 of MT, we write $D_i - K_i = H_1^i + H_2^i + \cdots + H_{s_i}^i$ $(i = 1, 2, \cdots, p)$. Let $R_1 = B_1 + \sum_{i=1}^{p} \sum_{j=1}^{s_1} H_j^i$. Since R_1 is an open set containing B_1, it follows that there is a positive number a such that R_1 contains $\overline{S_G(B_1, a)}$.

Now let W_2 denote that manifold-with-boundary which is a subset of M and whose boundary curves are the 1-spheres Γ_j^i which form part of the boundaries of the cylinders H_j^i $(i = 1, 2, \cdots, p; j = 1, 2, \cdots, s_i)$. To each curve Γ_j^i we attach a 2-cell E_j^i, thus obtaining the manifold W_2' (see §4). Now $F_j^i = E_j^i + H_j^i$ is a plane. We define an upper semi-continuous collection X_j^i filling F_j^i as follows:

(1) the elements of G which belong to $\overline{S_G(B_1, a)}$ are elements of X_j^i; and

(2) the remaining points of F_j^i are elements of X_j^i.

Then it follows from results of R. L. Moore[5] that X_j^i is a plane. Hence it follows that there is a simple closed curve C_{ij} in that part of X_j^i which arose from $S_G(B_1, a)$ such that C_{ij} separates the cylinder H_j^i between Γ_j^i and B_1. Thus C_{ij} is a simple closed curve both in the space X_j^i and in the space G. Moreover, there is a simple closed curve d_{ij} in the space M which separates the cylinder H_j^i between B_1 and C_{ij}. We define W_1 as that manifold-with-boundary which is a subset of M and whose boundary curves are the 1-spheres d_{ij}. Then W_1 and W_2 are homeomorphic and condition (6) of Lemma A is satisfied.

Now we define G_0 as that subset L of G which fills the integral compact component of $M - \sum_{i=1}^{p} \sum_{j=1}^{s_1} c_{ij}$, plus the boundary of L; that is, $G_0 = L$ plus the 1-spheres c_{ij}. We proceed to prove that G_0 is a manifold-with-boundary and that G_0 is homeomorphic to W_2. Let V_0 denote an upper semi-continuous collection filling W_2' (see above) and defined as follows:

(1) every element of G_0 is an element of V_0; and

(2) for each cylinder H_j^i, we define the elements of V_0 filling the plane F_j^i of which H_j^i is a subset as the elements of the collection X_j^i which fills F_j^i.

Then if g is an element of V_0, $R(g) = 0$. Hence by Theorem 1 of MT, V_0 is homeomorphic to W_2'. But G_0 and W_2 are obtained from V_0 and W_2' respectively by deleting a finite number (p in each case) of non-overlapping open 2-cells whose boundaries are 1-spheres. Hence G_0 is homeomorphic to W_2, and W_2 contains G_0 and G_0 contains W_1.

Now let n_1 be the smallest integer such that $M_{n_1}^0$ contains W_2. We let $W_3 = M_{n_1}$. Let A_2 denote the set of those elements of G each of which intersects the boundary of M_{n_1+1}, and let B_2 denote A_2 plus those elements of G

[5] R. L. Moore, *Concerning upper semi-continuous collections of continua*, Transactions of the American Mathematical Society, vol. 27(1925), pp. 416–428.

each of which lies in a compact component of $M - A_2$. Then we define W_4, W_5, and G_1 relative to B_2 just as we defined W_1, W_2 and G_0 relative to B_1, and we prove that G_1 is homeomorphic to W_4 and W_5 just as we proved that G_0 is homeomorphic to W_1 and W_2. In a similar manner we can prove that if T_1 is a component of $W_5 - W_1$, then that component of $G_1 - G_0$ which is contained in T_1 is homeomorphic to T_1. Moreover, we can continue this process by induction in such a way that all the conditions of Lemma A will be satisfied. Hence Theorem 1 is proved.

6. THEOREM 2. *If, for each g of G, $M - g$ is connected and has just one cylinder of approach to g* (see MT, Lemma 1), *then G is a 2-manifold without boundary, $R(G) \leqq R(M)$, and if M is orientable, so is G.*

Proof. The property of being a 2-manifold without boundary is a local property. Hence we need only repeat that, given any element g of G, there exists an integer n_g such that M_{n_g} contains g in its interior. It then follows at once from Theorem 2 of MT that g has a neighborhood in G which is a 2-cell. Therefore G is a 2-manifold without boundary.

If $R(M)$ is infinite, then surely $R(G) \leqq R(M)$. Suppose $R(M)$ is finite and $R(M) < R(G)$. Decompose G according to Lemma 1, $G = G_1 + G_2 + \cdots$. Then there exists an integer n_0 such that $R(M) < R(G'_{n_0})$ (see §4 for the meaning of G'_{n_0}). Now there exists an integer n_1 such that G_{n_0} arises entirely from elements of G in the interior of M_{n_0}. From this and from Theorem 2 of MT it follows that $R(G'_{n_0}) \leqq R(M'_{n_1})$. But surely $R(M'_{n_1}) \leqq R(M)$. These inequalities give a contradiction, and it follows that $R(G) \leqq R(M)$.

Now suppose M were orientable and G were not. Then we could find a 1-sphere d on G such that a directrix going around d would come back with the opposite orientation. Decompose G according to Lemma 1, $G = G_1 + G_2 + \cdots$. There exists an integer i such that M_i contains in its interior every element g of G_{n_d} (see §4). Let G_0 be an upper semi-continuous collection filling M_{n_i} defined as follows: given a point p of M_{n_i}, the element g of G_0 which contains p is defined as the element g_p of G which contains p or as the point p itself according as g_p intersects M_i or not. M_{n_i}, as a subset of M, is certainly orientable and hence G_0 is orientable, by Theorem 2 of MT, and the observation in footnote 10 and the sentence to which it refers in MT. But G_0 contains d and so is not orientable. This contradiction proves that G is orientable, and the theorem is complete.

7. Let g be an arbitrary compact continuum on M and let $g' = g$ plus all 2-cell complementary domains of g. Given any positive number e, there exists a domain V_e containing g' such that $S(g', e)$ contains V_e and the boundary of V_e is a finite set of mutually exclusive 1-spheres. To prove this statement, we note that if it is true for compact manifolds, then it follows readily for a non-compact manifold M, since g' is surely compact. We indicate a proof for a compact manifold M, using Lemma 1 of MT. There are only a finite number

of components of $M - g'$, say $D_1, D_2, \cdots, D_p$. Using the same modification of the notation of Lemma 1 of MT that we used in the proof of Theorem 1, we write $D_i - K_i = H_1^i + H_2^i + \cdots + H_{s_i}^i$ $(i = 1, 2, \cdots, p)$. Now we can find a 1-sphere on every H_j^i which is isotopic on M to Γ_j^i and lies in $S(g', e)$. This set of 1-spheres bounds the domain V_e whose existence was asserted.

Now let $a(g)$ be the smallest integer such that for every positive number e there exists a V_e whose boundary consists of $a(g)$ or fewer mutually exclusive 1-spheres. We call $a(g)$ the *order* of g. That $a(g)$ is finite for every g follows from the argument above, and in fact $a(g) = s_1 + s_2 + \cdots + s_p$. From the same argument we obtain the following

LEMMA 3. *Given an element g of G, there exists a positive number $e = e(g)$ such that if V_e and U_e are any two open sets each of which is contained in $S(g', e)$ and contains g' and is bounded by $a(g)$ or fewer mutually exclusive* 1*-spheres, then the* 1*-spheres bounding V_e and those bounding U_e are both exactly $a(g)$ in number, and there is a unique way of pairing them off with the following properties*:

(1) *each pair consists of just one* 1*-sphere from the boundary of each of the open sets; and*

(2) *the two* 1*-spheres of each pair are isotopic on $M \cdot S(g', e)$.*

The number of elements of G which are of order greater than 2 is finite on any M_i. In fact, if this were false for $i = i_0$, it would follow from part of the argument in the proof of the first part of Theorem 5 of MT that $R(M'_{n_{i_0}})$ was not finite.

DEFINITION. A *d-manifold* is a manifold different from a 2-sphere.

DEFINITION. A continuous curve[6] S is called an *A-space* if it has the following properties:

(1) every maximal cyclic element of S is a 2-manifold;

(2) if $J_1, J_2, \cdots$ denote the maximal cyclic elements of S which are d-manifolds ($J_i \neq J_k$ for $i \neq k$), then $\text{Ls}\, J_n = 0$;[7] and

(3) if p is any point of S, then there are only a finite number of the components of $S - p$ which are non-compact in S.

8. THEOREM 3. *If, for each element g of G, each component of $M - g$ has just one cylinder of approach to g, then G is an A-space, and $R(G) \leqq R(M)$.*

Proof. That G is a continuous curve follows from a theorem of Roberts.[8] But with our restrictive hypothesis, this statement can be proved directly with little trouble.

Now suppose that g is an element of G which belongs to a maximal cyclic

[6] A continuous curve is a space which is separable, metric, connected, locally connected, and locally compact.

[7] Ls J_n means the limit superior of the sequence $J_1, J_2, \cdots$. See C. Kuratowski, *Topologie* I, Warszawa-Lwów, 1933, p. 153.

[8] J. H. Roberts, *Concerning metric collections of continua*, American Journal of Mathematics, vol. 53(1931), p. 423, Theorem 1.

element X of S. We want to show that g has a 2-cell neighborhood in X. Let U be a compact open set in M containing g. Then there is a compact open set V of G such that $M_{n_g} \cdot U$ contains $\bar{V}$ and V contains g. We now define an upper semi-continuous collection G_0 filling M'_{n_g} as follows: given a point p of M'_{n_g}, the element of G_0 which contains p is defined as the element g_p of G which contains p, or as the point p itself according as g_p is an element of $\bar{V}$ or not. Then by Theorem 3 of MT, G_0 is a generalized cactoid. But G and G_0 are obviously locally homeomorphic at g since V is an open set about g both in G and in G_0. Hence every element of X has a neighborhood in X which is a 2-cell, and it follows that X is a 2-manifold.

We observe that the condition (2) in the definition of A-space is a consequence of the following

PROPOSITION 2′. *If p is a point of S, then there exists an open set U of S such that U contains p and only a finite number of maximal cyclic elements of S which are d-manifolds intersect U.*

We proceed to prove Proposition 2′. Let g be any element of G and let $g' = g$ plus all 2-cell complementary domains of g; g' is a compact continuum. Let V_e be an open set of M containing g', with the properties of the V_e of Lemma 3. It is an immediate consequence of the definition of an upper semi-continuous collection that the elements of G lying entirely inside an open set V of M form an open set of G. Let $b(V)$ denote the open set of G thus associated with the open set V of M. Then $b(V_e)$ is the open set U of G whose existence we are trying to prove. For, in the first place, if Y denotes the set of maximal cyclic elements of G each of which has the following properties:

(1) it is a d-manifold,

(2) it intersects $b(V_e)$, and

(3) it arises from elements of G belonging to $M - g'$;

then it follows from Lemma 3 that the set Y can contain at most one maximal cyclic element for each of the $a(g)$ 1-spheres which form the boundary of V_e. And secondly, by R. L. Moore's results[9] we know that any maximal cyclic element of G arising from a 2-cell complementary domain of g must be a 2-sphere. Hence Proposition 2′ is proved.

The proof that G satisfies condition (3) of the definition of an A-space is highly analogous to the proof of Proposition 2′. Finally, the proof that $R(G) \leqq R(M)$ is similar to the proof of that inequality in Theorem 2.

9. THEOREM 4. *Every A-space is the monotone image of a 2-manifold.*

Proof. Let $J_1, J_2, \cdots$ be those maximal cyclic elements of the A-space S which are d-manifolds, as in the definition of an A-space, and let P_i^j be the cut points of S which belong to J_i ($i, j = 1, 2, \cdots$). Let W_{ij}^a be the components of $S - P_i^j$ which are compact cactoids, B_{ij} the component of $S - P_i^j$ which

[9] R. L. Moore, *Concerning upper semi-continuous collections*, Monatshefte für Mathematik und Physik, vol. 36(1929), pp. 81–88.

contains $J_i - P_i^j$, X_{ij}^b the components which contain d-manifolds $\neq J_i$, and Y_{ij}^c the remaining components, where the range of a, b, and c is the set of positive integers (we shall prove presently that the ranges of b and c are finite for fixed i and j). Let $W_{ij} = P_i^j + W_{ij}^1 + W_{ij}^2 + \cdots$, $X_{ij} = P_i^j + X_{ij}^1 + X_{ij}^2 + \cdots$, and $Y_{ij} = P_i^j + Y_{ij}^1 + Y_{ij}^2 + \cdots$.

W_{ij} is a compact cactoid. For if U is a compact open set about P_i^j, it follows from condition (2) of the definition of an A-space and the local connectedness and local compactness of S that only a finite number of the sets W_{ij}^a have points not in U. Hence W_{ij} is compact and the fact that W_{ij} is a cactoid follows immediately from the definition of an A-space.

There is only a finite number of the sets X_{ij}^b for any given pair i, j. For suppose there were infinitely many. Then from condition (2) of the definition of an A-space and the fact that S is locally compact, it follows that the diameters of the sets X_{ij}^b for fixed i and j are bounded away from zero, say are all $> 3e > 0$. Hence each set $X_{ij}^b + P_i^j$ contains an arc c_b from P_i^j to some point q_b not in $S(P_i^j, e)$. Let U be a compact open set such that $S(P_i^j, e)$ contains U and U contains P_i^j. Let $2d$ be the distance from P_i^j to $S - U$, and let x_b be a point of c_b such that the distance from P_i^j to x_b is d ($b = 1, 2, \cdots$). Since U is compact, the points x_b have a limit point x. But then it can be shown that S is not locally connected at x. This contradiction proves that there is only a finite number, say b_{ij}, of the sets X_{ij}^b.

Now each of the sets $Y_{ij}^c + P_i^j$ is non-compact; for since it does not contain a d-manifold it is a cactoid, and would be a W_{ij}^a if it were compact. Hence by condition (3) in the definition of an A-space, there is only a finite number, say c_{ij}, of the sets Y_{ij}^c for any fixed pair i, j.

Now if F is a non-compact cactoid, then we can write $F = F_1 + F_2 + \cdots$, where, for every n, F_n is a compact cactoid and $F_n \cdot F_{n+1}$ is exactly one point, while $F_n \cdot F_p = 0$ for $p > n + 1$. If we use this fact and the results of Roberts and Steenrod, it follows that every non-compact cactoid is the monotone image of a 2-manifold.

Now let K_i be a replica of J_i, and choose 2-cells E_i^j on K_i, in accordance with Lemma 3 of MT, corresponding to P_i^j on J_i. Let $t = b_{ij} + c_{ij} + 1$ or $b_{ij} + c_{ij}$ according as W_{ij} contains a point $\neq P_i^j$ or not. Remove from the interior of E_i^j t non-overlapping open 2-cells whose boundaries are mutually exclusive 1-spheres. Let $N_1, N_2, \cdots, N_{c_{ij}}$ be 2-manifolds-with-boundary, the boundary of each of which is a single 1-sphere, and such that there is a monotone transformation which carries N_c into Y_{ij}^c ($c = 1, 2, \cdots, c_{ij}$). Likewise, let $L_1, L_2, \cdots, L_{b_{ij}}$ be closed and compact cylinders. Affix the bounding curve of each of the sets N_c to the bounding curve of one of the 2-cells removed from E_i^j, and affix one of the bounding curves of each of the cylinders L_b to the bounding curve of one of these 2-cells. If there is still a 2-cell left, that is, if $t = b_{ij} + c_{ij} + 1$, we affix to the boundary of this 2-cell the boundary of a closed 2-cell, that is, we replace the deleted 2-cell.

Now suppose J_s is a d-manifold in $X_{ij}^{b_1}$ and Q is a point of J_s. There is an

arc c from Q to P_i^j. Since any compact subset of S intersects only a finite number of the J_i by condition (2) of the definition, it follows that there is a last J_i which c intersects before reaching P_i^j. Call this last one J_q and let P_q^l be the last point of J_q on c, where the order of points on c is from Q first to P_i^j last. P_q^l depends only on the choice of J_i (the J_i containing P_i^j, that is) and J_q. For otherwise $X_{ij}^{b_1}$ would not be a maximal cyclic element of S. Clearly P_q^b is a cut point of J_q. When we treat J_q as we have been treating J_i heretofore in the proof, we shall find a set $X_{ql}^{b_2}$ which contains J_i, and to it will correspond the boundary of an open 2-cell which will have been removed from the interior of the 2-cell E_q^l. It is to this boundary that we affix the 1-sphere which forms the boundary of the other end of the cylinder L_{b_1} (recall that b_1 is that value of b such that $X_{ij}^{b_1}$ contains J_s).

We go through this procedure first for every P_1^j ($j = 1, 2, \cdots$), then for every $P_2^j, P_3^j, \cdots$. The result is a manifold M which has been obtained from a sequence of manifolds $K_1, K_2, \cdots$ by attaching manifolds and by joining pairs of the K_i with cylinders.

The monotone transformation T is defined over M as follows: $T(K_i - (E_i^1 + E_i^2 + \cdots))$ is homeomorphic to $J_i - (P_i^1 + P_i^2 + \cdots)$ ($i = 1, 2, \cdots$); let V_i^j denote E_i^j minus the 2-cells deleted from E_i^j; then $T(V_i^j) = P_i^j$ ($i, j = 1, 2, \cdots$); the image under T of the cylinder L_{b_1} between E_i^j and E_q^l corresponding to P_i^j and P_q^l is the common part of the two sets $(X_{ij}^{b_1} + P_i^j)$ and $(X_{ql}^{b_2} + P_q^l)$, and this common part is a compact cactoid. The image under T of the manifolds N_c attached to E_i^j are the non-compact cactoids Y_{ij}^c. The image under T of the 2-cell attached to E_i^j is the cactoid X_{ij}. Thus it follows that M and T are the manifold and the monotone transformation whose existence the theorem asserts.

10. Definition. An *A-space with identifications* is a space which is obtained from an A-space S as follows: points P_j^i of S are chosen ($i = 1, 2, \cdots; j = 1, 2, \cdots, k_i$) in such a way that there is no point of S which is a limit point of these points, and then for every i the set of points P_j^i ($j = 1, 2, \cdots, k_i$) are identified.

Theorem 5. *For any manifold M and any collection G filling M, G is an A-space with identifications and $R(G) \leqq R(M)$; and conversely, for every A-space with identifications K, there exist an M and G filling M such that G is homeomorphic to K.*

We give only a brief indication of the proof.

The first part of this theorem is proved by the method used to prove the first part of Theorem 5 of MT. In general it will be necessary to make a cut around each of a countable infinity of 1-spheres and to affix a 2-cell onto each of the lips of each of the 1-spheres which was cut around. But there will be only a finite number of these cuts necessary on any M_i, and hence the restriction that identified points have no limit point is satisfied.

The proof of the converse part of the theorem is the same as the proof of Theorem 4 except for a slight modification. In order to obtain the identifications we introduce 2-cells corresponding to the points to be identified just as we introduced the 2-cells E_i^j, and we introduce cylinders joining the 2-cells; that is, if we have a finite set of points to be identified, we introduce a cylinder joining every pair of the corresponding set of 2-cells. Then all these cylinders and the 2-cells they join will be a single element of the new upper semi-continuous collection.

Part II

11. Roberts has shown by an example[10] that the space whose elements are the continua of an upper semi-continuous collection filling a 2-manifold, not all of which are bounded, need not be metric. In this paper, however, we restrict the class of spaces G by assuming that G is metric. We shall say $M = M_1 + M_2 + \cdots$, meaning always that this decomposition is in accordance with Lemma 1 of Part I.

12. Lemma 1. *If g is an element of G and a is a positive number, then there exists an integer $n = n(g, a)$ such that $n^{-1} < a$ and such that if h is an element for which $\rho(h, g) \leqq n^{-1}$, then $h \cdot M_n^0 \neq 0$.*

Proof. Suppose the lemma is false. Then for every integer k there is an element g_k such that $\rho(g, g_k) \leqq k^{-1}$ and $g_k \cdot M_k^0 = 0$. Then g is the limit element of the sequence $g_1, g_2, \cdots$. But if p is a point of g, then for some integer k_0, p belongs to $M_{k_0}^0$; and since only a finite number of the elements of the sequence $g_1, g_2, \cdots$ intersect $M_{k_0}^0$, p cannot be a limit point of the point set $g_1 + g_2 + \cdots$. Hence no point of g is a limit point of $g_1 + g_2 + \cdots$, and it follows from the definition of limit element (see CC) that g is not a limit element of the sequence $g_1, g_2, \cdots$. This contradiction proves the lemma.

Lemma 2. *If $c_1, c_2, \cdots$ is a fundamental sequence of cuts,[11] and if g is an element of G which intersects c_k for infinitely many values of k, and if no other element of G intersects c_k for infinitely many values of k, then there is an integer n such that every element which intersects that component K_n of $M - c_n$ which contains c_{n+1} also intersects c_n.*

Proof. We remark first that it follows from Lemma 1 of Part I and the definition of a fundamental sequence of cuts that there is a fundamental sequence of cuts $d_1, d_2, \cdots$ which is equivalent to the sequence $c_1, c_2, \cdots$ and which has the property that, for every k, d_k is a boundary curve of M_k. If we can prove that there is an integer n such that every element which intersects that component J_n of $M - M_n$ which contains d_{n+1} also intersects d_n, then the lemma will be proved.

[10] J. H. Roberts, *Concerning collections of continua not all bounded*, American Journal of Mathematics, vol. 52(1930), p. 552. The definitions there given of upper semi-continuous collection and limit element are the definitions we employ. This paper will be referred to hereafter as CC.

[11] See Kerékjártó, loc. cit., p. 164.

For every integer s we define a set H_s of G as follows: an element h belongs to H_s if $h \cdot d_s \neq 0$ and $\rho(h, g) \leqq s^{-1}$. Now the set of all those elements each of which intersects a closed and compact set in M is closed and compact in G Hence for every s, H_s is closed and compact.

We now prove the following

ASSERTION. *If A_s denotes the set of those elements of $G - H_s$ each of which lies entirely in J_s, then there exists an integer s_1 such that $A_s = 0$ for $s > s_1$.*

Proof. First we prove that if $s_1 > s$, then A_s contains A_{s_1}. Let h belong to A_{s_1}. Then h lies entirely in J_{s_1} and hence *a fortiori* in J_s. It follows then that $h \cdot d_s = 0$. Consequently h cannot belong to H_s. Hence by definition of A_s, h belongs to A_s.

Now let $\frac{1}{8}$ be the number a of Lemma 1. Then there exists an integer k such that $k > 8$ and such that if h is an element for which $\rho(h, g) \leqq k^{-1}$, then $h \cdot M_k^0 \neq 0$. From this it follows that $\rho(g, A_k) \geqq k^{-1}$, and in fact that $\rho(g, A_s) \geqq k^{-1}$ for $s \geqq k$.

Now let B denote that set of elements each of which intersects d_{k+1} but whose distance from g in the space G is not less than $(2k)^{-1}$. B is closed and compact. Now there exists an integer $v > k$ such that $B \cdot J_v = 0$. For suppose the contrary; let h_m be an element of B such that $h_m \cdot J_m \neq 0$ ($m = k + 2, k + 3, \cdots$), and such that the sequence $\{h_m\}$ converges to the element h_0. By the definitions of upper semi-continuous collection and limit element it follows that $h_0 \cdot d_{k+1} \neq 0$. By the hypothesis of the lemma there is an integer $b > k + 1$ such that $h_0 \cdot d_b = 0$. But infinitely many of the elements h_m intersect d_b and hence by the definitions of upper semi-continuous collection and limit element it follows that $h_0 \cdot d_b \neq 0$. This contradiction proves the existence of the integer v.

Now I say that v is the integer s_1 whose existence we are trying to establish; that is, that $A_v = 0$. For suppose h is an element of A_v. Let pq be an arc in J_v from a point p of h to a point q of g. Now $\rho(g, h) \geqq k^{-1}$, for we have shown above that $\rho(g, A_s) \geqq k^{-1}$ for $s \geqq k$, and we chose $v > k$. Since the set of those elements each of which intersects the connected set pq is connected, it follows that there is an element g_1 which intersects pq and which has the property that $(2k)^{-1} < \rho(g, g_1) < k^{-1}$. Since $g_1 \cdot J_v \neq 0$, it follows that g_1 does not belong to the set B, and hence that $g_1 \cdot d_{k+1} = 0$. Hence g_1 lies entirely in J_{k+1}, and g_1 does not belong to H_{k+1}. Then by the definition of A_s it follows that g_1 belongs to A_{k+1}. But $\rho(g, g_1) < k^{-1}$, and we saw above that $\rho(g, A_s) \geqq k^{-1}$ for $s \geqq k$. This contradiction proves the assertion.

Now I say that s_1 is the integer n whose existence the lemma asserts. For otherwise there is an element h which intersects J_{s_1} but does not intersect d_{s_1}. Then h lies entirely in J_{s_1} and h does not belong to H_{s_1}. Hence h belongs to A_{s_1}. But we have shown that A_{s_1} is vacuous, and this contradiction proves the lemma.

LEMMA 3. *No compact element of G is a limit element of non-compact elements.*

Proof. Suppose the lemma is false and let g be a compact element which is the sequential limit element of a sequence of non-compact elements $h_1, h_2, \cdots$. Now let U be a compact open set in M containing g. Then no h_i lies entirely in $\bar{U}$. Hence there is a point p of $\bar{U} - U$ which is a limit point of $h_1 + h_2 + \cdots$ but which does not belong to $h_1 + h_2 + \cdots$. Then by the definition of limit element, the element g_p which contains p is a limit element of $h_1 + h_2 + \cdots$. Since g is the sequential limit element of $h_1 + h_2 + \cdots$, this means that $g_p = g$. But U contains g and hence g does not contain p. This contradiction proves the lemma.

13. Now if $R(M)$ is finite, there is a compact manifold F such that M is homeomorphic to $F - (p_1 + p_2 + \cdots + p_k)$ where p_i is a point of F ($i = 1, 2, \cdots, k$). Throughout the rest of the paper we shall consider only the case in which $R(M)$ is finite, and we shall make the assumption (which, in view of the remark above, adds no essential restriction) that $M = F - (p_1 + p_2 + \cdots + p_k)$. We shall find it convenient to speak of "an element g having p_i as a limit point", meaning of course that p_i is a limit point in F of the set of points in F each of which belongs to the element g. Likewise we shall speak of the sphere $S_M(p_i, a)$, of the distance $\sigma(p_i, q)$, etc.

LEMMA 4. *Suppose there is a single element g which has p_i as limit point. Let G_1 be the collection of continua filling $M + p_i$ whose elements are the same as the elements of G except that g is replaced by $g + p_i$. Then G_1 is upper semi-continuous and metric, and G_1 is homeomorphic to G.*

Proof. We need only to show that the element $g + p_i$ of G_1 is a limit element of a sequence of elements only if g is a limit element of the same sequence. Let $h_1, h_2, \cdots$ be a sequence of elements having the sequential limit element $g + p_i$. Then by Lemma 2 there is an n such that, for every i, h_i intersects the boundary of M_n. From this it follows by the theorem of Zoretti that Ls h_i (see footnote 7) is connected. Then the element h of G which is the limit element in G of the sequence $h_1, h_2, \cdots$ must contain a point in every neighborhood of p_i. Hence $h = g$ and the lemma is proved.

In view of Lemma 4 we are able to assume that there is not exactly one element which has p_i as limit point. For if there is, we add p_i to M and to that element which has p_i as limit point, and G is not affected thereby. We shall make this assumption throughout the rest of the paper. The justification of this assumption has been the chief purpose of Lemmas 1, 2, and 4.

14. LEMMA 5. *If g is a continuum on a compact manifold F, and if $F - g$ is connected, then a necessary and sufficient condition that $R(g) = 0$ is that there exist a closed 2-cell in F having g in its interior.*

The necessity of the condition is given by Lemma 2 of MT. The proof of the sufficiency follows from the fact that a continuum with positive Betti number lying in an open 2-cell separates that 2-cell.

If g is an element of G, we shall denote by g' the closure in F of that set of points each of which belongs to the element g plus all the 2-cell complementary domains of that closed point set.

THEOREM 1. *Let F be a compact manifold such that $M = F - (p_1 + p_2 + \cdots + p_k)$, where p_i is a point of F $(i = 1, 2, \cdots, k)$. Suppose that, for every non-compact element g of G, $R(g') = 0$, and $F - g'$ is connected and has a single cylinder of approach to g'.* (See Lemma 1 of MT.) *Suppose also that no non-compact element of G has both p_i and p_j as limit point for $i \neq j$. Then if Q_i is the point set of F covered by those non-compact elements of G each of which has p_i as limit point, it follows that there is a closed 2-cell L_i in F whose boundary is a simple closed curve and such that L_i contains in its interior $Q_i + p_i$ and such that $L_i \cdot L_j = 0$ for $i \neq j$ $(i, j = 1, 2, \cdots, k)$.*

Proof. If p is a point of M and h_1 and h_2 are two non-compact elements such that h_1' contains p and h_2' contains p, then either h_1' contains h_2' or else h_2' contains h_1'. If $h_1, h_2, \cdots$ is a sequence of non-compact elements such that for $n = 1, 2, \cdots$ h_{n+1}' contains $h_n' + p$, then there is a limit element h of this sequence which has the property that h' contains $h_1' + h_2' + \cdots$, and h is non-compact. This follows readily from Lemma 3 and the theorem of Zoretti. Hence there is a non-compact element h_p such that h_p' contains p, and if h is a non-compact element for which h' contains p, then h_p' contains h'.

Now we define a new collection G_1 filling M as follows:

(1) if p is a point for which there is a non-compact element h such that h' contains p, and if h_p' is as above, then $h_p' - (p_1 + p_2 + \cdots + p_k)$ is an element of G_1, and

(2) if p is a point for which there is no non-compact element h such that h' contains p, then the element of G_1 which contains p is the same as the element of G which contains p.

It is easily verified that G_1 is upper semi-continuous and metric, that no non-compact element of G_1 separates M or F, and that if we can find 2-cells in accordance with the theorem which contain in their interiors the non-compact elements of G_1, then these 2-cells will contain in their interiors the non-compact elements of G. Hence we proceed to prove the theorem for the collection G_1 instead of for the collection G.

If the theorem is true for the case $k = 1$, it follows readily for every integer k. For the distance from Q_i to Q_j is positive for $i \neq j$; and since we can close down on the set $Q_i + p_i$ by a sequence of simple closed curves bounding 2-cells, it follows that if we choose simple closed curves which are far enough along in these sequences, then $L_i \cdot L_j$ will be vacuous for $i \neq j$.

Let us consider the case $k = 1$ then, and let $p_1 = p$, $Q_1 = Q$. Let c be a 1-sphere in F bounding a 2-cell E which contains p in its interior, and let g_0 be any non-compact element of G which intersects c. Let q_0 be a point of $c \cdot g_0$, and let f be a homeomorphism which carries the circle $\rho = 1$ of the Euclidean plane into c in such a way that $f(1, 0) = q_0$; let $f(1, \theta)$ be denoted by q_θ

$(0 \leqq \theta \leqq 2\pi)$. Then if g is a non-compact element such that $g \cdot c$ contains q_b, then g_b will denote the element g. This admits the possibility that $g_a = g_b$ even though $a \neq b$. If b is any positive number such that $b \leqq 2\pi$, then we let K_b denote the point set of F covered by p plus every non-compact element that may be denoted by g_x for some number $0 \leqq x \leqq b$. Now suppose that T is a component of $E - K_b$ and suppose every non-compact element which intersects T lies entirely in T. Let D_T denote the point set in T covered by non-compact elements. We define Q_b as the set K_b plus every such set D_T which has the property that K_b contains $\bar{D}_T - D_T$. Then Q_b is a compact continuum in F, and $Q_{2\pi} = Q$.

Now if $R(Q) = 0$ it follows by Lemma 5 that our theorem is true. Hence suppose $R(Q) \neq 0$.

We now prove

Lemma A. *If W is any closed set of non-compact continua such that there exists a 2-cell A in F which contains $W + p$ in its interior, then $W + p$ does not separate F.*

Proof. Let t be a homeomorphism which maps A into a 2-sphere S. Now if $t(W + p)$ does not separate S, it follows that $W + p$ does not separate F. But $S - t(p)$ is a plane and $t(W - p)$ is a closed point set in $S - t(p)$ which is a sum of mutually exclusive unbounded continua no one of which separates $S - t(p)$. Hence by Theorem 1 of CC, $t(W - p)$ does not separate $S - t(p)$. Then $W + p$ does not separate F, and the lemma is proved.

Now let X denote the set of non-negative real numbers which have the property that if x is in X, then there is a closed 2-cell A_x which contains Q_x in its interior. Let Y denote the set of those positive real numbers each of which is less than or equal to 2π and does not belong to X. Since G is upper semi-continuous, it follows readily that X does not contain a largest number.

Now if Y is vacuous then there is a closed 2-cell which contains $Q_{2\pi} = Q$ in its interior, and the theorem follows. Hence suppose that Y contains a smallest number y_0. Then there is a non-compact element g_{y_0} and a closed 2-cell H which contains g'_{y_0} in its interior. Moreover, there is a positive number v such that if V denotes the set of all those non-compact elements h for which $\rho(h, g_{y_0}) \leqq v$, then H contains $V + p$ in its interior. Now let w be a non-negative number less than y_0 such that

(1) if there is a largest number w less than y_0 for which g_w has a meaning, then w is this largest number, and

(2) if there is no such largest number, then w is a number for which g_w has a meaning and which is such that $y_0 - w$ is so small that if $y_0 > d \geqq w$ and g_d is defined, then g_d belongs to V. Let V_1 denote the closure of $V - Q_w$.

Now by Lemma A, neither $V_1 + p$ nor Q_w separates F. Hence by Lemma 5, $R(V_1 + p) = R(Q_w) = 0$. Then by a theorem of Borsuk[12] the space of con-

[12] K. Borsuk, *Über die Abbildungen der metrischen kompakten Räume auf die Kreislinie*, Fundamenta Mathematicae, vol. 20(1933), Theorem H′, p. 230.

tinuous mappings of $V_1 + p$ into a 1-sphere is connected, as is the space of mappings of Q_w into a 1-sphere. Also the set $(V_1 + p)\cdot Q_w$ is connected since it contains p and is an integral set. Hence by a theorem of Eilenberg[13] the space of mappings of $V_1 + p + Q_w$ into a 1-sphere is connected. Then by Borsuk's theorem, $R(V_1 + p + Q_w) = 0$.

Now let Z denote the closure of $Q_{y_0} - (Q_w + V_1 + p)$. Then by Lemma A, Z does not separate F; also the 2-cell E contains Z in its interior. Hence by Lemma 5, $R(Z) = 0$. But now as before $Z\cdot(Q_w + V_1 + p)$ is connected and it follows that $R(Z + Q_w + V_1 + p) = 0$. But $Z + Q_w + V_1 + p = Q_{y_0}$. By Lemma A again, Q_{y_0} does not separate F, and so by Lemma 5 there is a closed 2-cell containing Q_{y_0} in its interior. This contradicts the definition of y_0 and proves the theorem.

15. Definition. By a *manifold with deletions* we mean a space which can be obtained from a compact 2-manifold in a finite number of steps, each step consisting of the deletion of either

(a) a single point, or

(b) an open 2-cell bounded by a 1-sphere, or

(c) a point p plus a set A of mutually exclusive open 2-cells each of which is bounded by a 1-sphere passing through p and such that the limit superior of the set of 1-spheres bounding 2-cells of the set A contains no point except possibly p.

We remark that a 2-sphere with a single deletion is either a plane or else the domain D plus boundary described in Theorem 4 of CC. The above definition was framed so as to permit the application of Theorem 4 of CC in the following

Theorem 2. *If the hypotheses of Theorem 1 are satisfied, then there exists a finite set of spaces $C_0, C_1, \cdots, C_s$ with the following properties:*

(1) *C_0 is a continuous curve whose every maximal cyclic element is either a 2-manifold or a 2-manifold with deletions,*

(2) *all but a finite number of the maximal cyclic elements of C_0 are either 2-spheres or 2-spheres with deletions,*

(3) *C_i is obtained from C_{i-1} by identifying just two points of C_{i-1} ($i = 1, 2, \cdots, s$), and*

(4) *C_s is homeomorphic to G.*

Proof. Let $L_1, L_2, \cdots, L_k$ be the closed 2-cells of Theorem 1 which contain in their interiors all the non-compact elements. Then it follows that there exist connected open sets V_i of G such that for $i = 1, 2, \cdots, k$

(1) L_i contains $\bar{V}_i$ in its interior,

(2) $V_1 + V_2 + \cdots + V_k$ contains all the non-compact elements,

(3) the distance from p_i to the boundary of $V_i + p_i$ is positive,

(4) $R(\bar{V}_i + p_i) = 0$, and

[13] S. Eilenberg, *Sur les transformations d'espaces métriques en circonférence,* Fundamenta Mathematicae, vol. 24(1935), Theorem 3, p. 162.

(5) $\bar{V}_i$ contains no element g of G such that $R(g') \neq 0$.

Now we define a new collection G_1 filling F as follows:

(1) $\bar{V}_i + p_i$ is the element g_i of G_i ($i = 1, 2, \cdots, k$) and

(2) every element of G not lying in $\bar{V}_1 + \bar{V}_2 + \cdots + \bar{V}_k$ is an element of G_1.[14]

Then by Theorem 5 of MT, G_1 is a space which can be obtained from a generalized cactoid K_1 by making a finite number of identifications. We may think of K_1 as being obtained from G_1 by restoring the points of G_1 which came from identifications of K_1 to their original multiplicity; that is, by deleting a finite number of local cut points of G_1 and then replacing each of these points multiplicately. Let K be the space obtained from G by deleting from G the same (see footnote 14) finite set of local cut points and replacing them multiplicately in the same way. Now since K_1 is a generalized cactoid, it follows that every maximal cyclic element of K which contains no point of $\bar{V}_1 + \bar{V}_2 + \cdots + \bar{V}_k$ is a 2-manifold, and only a finite number of these 2-manifolds are not 2-spheres. Hence, since G can be obtained from K by making a finite number of identifications, if we can prove the following lemma it follows that our theorem is proved (where K is the space C_0 of the theorem).

Lemma B. *Every maximal cyclic element of K which contains a point of $\bar{V}_i$ is a 2-manifold or a 2-manifold with deletions, and all except possibly one of these maximal cyclic elements is a 2-sphere or a 2-sphere with deletions* ($i = 1, 2, \cdots, k$).

Proof. *Case* I. Let A be a maximal cyclic element of K which is contained in $\bar{V}_i$. Then we define a plane S and an upper semi-continuous collection of continua H filling S as follows: S consists of $L_i - p_i$ plus a 2-cell bounded by a 1-sphere, the 1-sphere being identified with the boundary of L_i by a homeomorphism; each element of $\bar{V}_i$ is an element of H; the remaining points of S are elements of H. Then A is a maximal cyclic element of H as well as of K. Then by a theorem of Roberts,[15] A is a 2-sphere or a 2-sphere with deletions.

Case II. Now let B be a maximal cyclic element of K which contains an element of $\bar{V}_i$ and also an element h which belongs to K_1 as well as to K (see footnote 14) and let B_1 be the maximal cyclic element of K_1 which contains h. Clearly B_1 contains g_i since B contains an element of $\bar{V}_i$. Now g_i is not one of the points of K_1 which is identified to give G_1. For points of identification of G_1 arise only from elements which separate the manifold locally, and which therefore have positive Betti number; and $g_i = \bar{V}_i + p_i$, and $R(\bar{V}_i + p_i) = 0$. Hence there is a neighborhood U of g_i in K_1 which contains no point of K_1 to be identified and contains no g_j for $j \neq i$, and such that every cyclic element

[14] Clearly $G - (\bar{V}_1 + \bar{V}_2 + \cdots + \bar{V}_k)$ is homeomorphic to $G_1 - (g_1 + g_2 + \cdots + g_k)$. We shall say that an element h of G is the same as the element h_1 of G_1, meaning that they cover the same point set in F.

[15] Theorem 2 of the paper referred to in footnote 8. The proof given there actually shows that each cyclic element is a 2-sphere or a 2-sphere with deletions, although the theorem states only that it is a subset of a 2-sphere.

of K_1 intersecting U, except possibly B_1, is a 2-sphere. Then $U - g_i$ is an open set in K such that the distance in K from $\bar{V}_i$ to the complement of $U + \bar{V}_i$ is positive. It follows then that B is the only maximal cyclic element of K which contains an element of $\bar{V}_i$ and an element which belongs both to K_1 and to K. Hence we have proved the last statement of the lemma and it remains only to prove that B is a 2-manifold or a 2-manifold with deletions. To prove this we remark first that if for every j such that $\bar{V}_j \cdot B \neq 0$ we can find an open 2-cell U_j in B_1 whose boundary is a 1-sphere and such that $\bar{U}_j$ is a 2-cell containing g_j but not containing g_v ($v \neq j$), then the closed set W_j in B which consists of the elements of $\bar{U}_j - g_j$ plus the elements of $\bar{V}_j \cdot B$ is a 2-cell with deletions. For we may imbed W_j in a plane exactly as we imbedded A in a plane in Case I and our result follows as it did in that case. But it is trivial that such 2-cells U_j can be found in B_1, since B_1 is a 2-manifold. Hence the lemma is proved and so is the theorem.

16. THEOREM 3. *If $M = F - (p_1 + p_2 + \cdots + p_k)$, then there exists a finite set $h_1, h_2, \cdots, h_m$ of the non-compact continua of G such that if M' is a component of $M - (h_1' + h_2' + \cdots + h_m')$ and G' is the subcollection of G which fills M', then M' and G' satisfy the hypotheses of Theorems* 1 *and* 2.

Proof. Let h_1 be an element of G which has as limit point r of the points p_i, where r is greater than one. Then if A_1 is a component of $M - h_1'$, A_1 is homeomorphic to a compact manifold minus at most $k - r + 1$ points instead of minus k points. It follows that there are v elements ($v \leqq k - 1$), $h_1, h_2, \cdots, h_v$, such that the following is true: if D is a component of $M - (h_1' + h_2' + \cdots + h_v')$, and if D is homeomorphic to $D_1 - (q_1 + q_2 + \cdots + q_s)$, where D_1 is a compact manifold and q_i is a point of D_1 ($i = 1, 2, \cdots, s$), and if h is a non-compact element lying in D, then h has as limit point only one of the points q_i. Moreover, $M - (h_1' + h_2' + \cdots + h_v')$ has only a finite number of components; for each of its components is a manifold different from a 2-cell, and $R(M)$ is finite. Now let A_2 denote a component of $M - (h_1' + h_2' + \cdots + h_v')$ and let A_2 be homeomorphic to a compact manifold F_2 minus a finite set of points. Let h_{v+1} be a non-compact element in F_2 for which $F_2 - h_{v+1}'$ is not connected. Then each of the components of $F_2 - h_{v+1}'$ is a manifold with smaller Betti number than F_2, but no such component is a 2-cell, for h_{v+1}' contains all 2-cell complementary domains of h_{v+1}. Hence we see that there is a finite set of elements $h_{v+1}, h_{v+2}, \cdots, h_{v+z}$ such that if A_3 is a component of $M - (h_1' + h_2' + \cdots + h_{v+z}')$ and h is a non-compact element lying in A_3, then $A_3 - h'$ is connected.

The proof that there is a finite set of elements $h_{v+z+1}, \cdots, h_{v+z+w}$ such that if A_4 is a component of $M - (h_1' + \cdots + h_{v+z+w}')$ and h is an element lying in A_4 then $A_4 - h'$ has a single cylinder of approach to h' follows from an argument given early in the proof of the first part of Theorem 5 of MT. Finally, it can be proved that there is a finite set of non-compact elements $h_{v+z+w+1}, \cdots, h_{v+z+w+y}$ such that if A_5 is a component of $M - (h_1' + \cdots + h_{v+z+w+y}')$ and h is

an element in A_5, then $R(h') = 0$. Thus our theorem follows when we place $m = v + x + w + y$.

THEOREM 4. *With only the two hypotheses that $R(M)$ is finite and that G is metric the conclusion of Theorem 2 holds.*

Proof. The proof proceeds by showing how the elements contained in the set $h_1' + \cdots + h_m'$ of Theorem 3 are fitted back into the space G. In the first place, if A is a maximal cyclic element of G arising entirely from elements of h_1', then it follows from a theorem of Roberts (see footnote 15) that A is a 2-sphere or a 2-sphere with deletions. Now h_i must be a limit element of $K = G - (h_1' + h_2' + \cdots + h_m')$ $(i = 1, 2, \cdots, m)$. But by a theorem already cited[16] G is locally compact, and also K is locally compact. Hence it follows from the description of K in Theorem 2 that h_i is either a limit point of one or more of the open curves arising from a deletion or else is the limit superior of a sequence of maximal cyclic elements of K. But the open curves of any component B of K arising from a deletion have a discrete cyclic order; and it can be proved that, since G is locally compact, if there is a single open curve c arising from a deletion and if c has h_i as limit element, then $c + h_i$ is a simple closed curve; whereas if there is more than one open curve arising from a deletion, and if c is one of these curves which has h_i as limit element, then there is another open curve d arising from this deletion which also has h_i as limit element in such a way that $c + h_i + d$ is a single open curve. It follows that $K + h_1 + h_2 + \cdots + h_m$ is a space which satisfies the first three properties in the conclusion of Theorem 2, since the additional identifications necessary as a result of adding $h_1 + h_2 + \cdots + h_m$ to K are clearly finite. And now our theorem follows since we have already shown that every maximal cyclic element of G which arises entirely from elements of h_i' is either a 2-sphere or a 2-sphere with deletions.

DUKE UNIVERSITY.

[16] The theorem referred to in footnote 8.

OPERATION THEORY AND MULTIPLE SEQUENCE TRANSFORMATIONS

By J. D. Hill and H. J. Hamilton

Consider the problem: To derive sets of conditions on the $(l + n)$-dimensional matrix of complex numbers $(a_{m^{(1)},m^{(2)},\dots,m^{(l)};k^{(1)},k^{(2)},\dots,k^{(n)}})$ $(l + n \geqq 2)$ necessary and sufficient that the l-tuple sequence $\{\sigma_{m^{(1)},m^{(2)},\dots,m^{(l)}}\}$ belong to a prescribed class whenever the n-tuple sequence of complex numbers $\{s_{k^{(1)},k^{(2)},\dots,k^{(n)}}\}$ belongs to a prescribed class, where

$$\sigma_{m^{(1)},m^{(2)},\dots,m^{(l)}} = \sum_{k^{(1)},k^{(2)},\dots,k^{(n)}=1}^{\infty} a_{m^{(1)},m^{(2)},\dots,m^{(l)};k^{(1)},k^{(2)},\dots,k^{(n)}} s_{k^{(1)},k^{(2)},\dots,k^{(n)}}.$$

This problem was solved in[1] H_1 and H_2 for each of the 256 cases corresponding to 16 classes of multiple sequences ranging from that in which the sequence is convergent and all partial limits exist and are zero to that in which the elements of the sequence are merely bounded for all values of the subscripts which are sufficiently large.

However, these derivations were based exclusively on "classical" methods, and this fact left open for a time the question of applicability of operation methods to multiple sequence transformations. The authors have now succeeded in applying these methods in the cases treated in H_1 and H_2 and our present purpose is to exhibit such phases of this application as may conceivably be of value in future investigations of multiple sequence transformations.

In case $l = n = 1$ we have to deal with the matrix (a_{mk}) and sequences $\{s_k\}$ and $\{\sigma_m\}$ related by the equation $\sigma_m = \sum_{k=1}^{\infty} a_{mk} s_k$. The classes of sequences which seem to have been of interest here are those of null, convergent, and bounded sequences.

We may remark parenthetically that the determination of "regularity" conditions on (a_{mk}), that is, conditions such that σ_m converge to $\lim_{k\to\infty} s_k$ whenever the latter exists, constitutes no separate problem, since we need here merely to determine conditions on the matrix (b_{mk}) necessary and sufficient that $\{\tau_m\}$ be a null sequence whenever $\{s_k\}$ is convergent, where $b_{mk} = a_{mk} - \delta_m^k$ and $\tau_m = \sum_{k=1}^{\infty} b_{mk} s_k$, and δ_m^k is Kronecker's symbol. Similar reduction to homogeneous

Received March 8, 1940.

[1] H_1 and H_2 denote the papers by Hamilton, *Transformations of multiple sequences*, this Journal, vol. 2(1936), pp. 29–60, and *Change of dimension in sequence transformations*, ibid., vol. 4(1938), pp. 341–342, respectively.

cases may be made when one considers the several regularity problems involved in multiple sequence transformations.[2]

The six transformations in which $\{s_k\}$ is either null or convergent have been adequately treated by operational methods in the literature, or at least can be so treated by simple extensions of the methods used in, say,[3] B, p. 91. If $\{s_k\}$ is merely bounded, the situation is more complicated, and, by way of showing how such difficulties as arise may be overcome, we treat the sufficiently typical case wherein $\{s_k\}$ is bounded and $\{\sigma_m\}$ is to be convergent.

The space of points $x = \{s_k\}$ under the norm $\| x \| = \sup_k | s_k |$ and the usual operational definitions is clearly a Banach space, and in it the points $x_p = \{\delta_k^p\}$ $(p = 1, 2, \cdots)$ and $x_{\{k_\nu\}}$ constitute a fundamental set,[4] where $x_{\{k_\nu\}}$ is the sequence consisting of 1's in the k_ν-th places and 0's elsewhere, $\{k_\nu\}$ being the general infinite[5] subsequence of the natural numbers k.

To prove this we fix upon x and take $R > \sup_k | s_k |$. Letting n be an arbitrary positive integer, we see that, corresponding to each k, there exist exactly one κ and one λ $(1 \leqq \kappa, \lambda \leqq n)$ such that

(1) $\quad R(\kappa - 1)/n \leqq | s_k | < R\kappa/n \quad \text{and} \quad 2\pi(\lambda - 1)/n \leqq \arg s_k < 2\pi\lambda/n.$

Let now $x_{\kappa,\lambda}$ be the sequence consisting of 1's for those values of k for which (1) is satisfied and 0's elsewhere. It follows that

$$\Big\| \sum_{\kappa,\lambda=1}^{n} [R(\kappa - 1)/n] \exp [2\pi i(\lambda - 1)/n] x_{\kappa,\lambda} - x \Big\|$$

can be made as small as desired by taking n sufficiently large, and this fact establishes the assertion.

Since $\sigma_m \equiv \sum_{k=1}^{\infty} a_{mk} s_k$ is to exist for each $\{s_k\} \epsilon (m)$, where (m) is the space of bounded sequences, we see by B, p. 86, that $\sum_{k=1}^{\infty} | a_{mk} | < \infty$. Hence the operation U_m defined by $U_m(x) = \sum_{k=1}^{\infty} a_{mk} s_k$ is linear, and it is easy to see that $\| U_m \| = \sum_{k=1}^{\infty} | a_{mk} |$. And since $U_m(x) \equiv \sigma_m$ is to converge for each $x \epsilon (m)$, it follows that $U(x) \equiv \lim_{m \to \infty} U_m(x)$ is linear, by B, p. 23, Theorem 4. Hence, by B, p. 80, Theorem 5 and p. 79, Theorem 3, the desired set of conditions may be written as follows:

[2] See Hamilton, *Preservation of partial limits in multiple sequence transformations*, this Journal, vol. 5(1939), pp. 293–297.

[3] By B we denote Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.

[4] B, p. 58.

[5] This unnecessary restriction is justified by its convenience.

$$\sup_m \sum_{k=1}^{\infty} | a_{mk} | < \infty ; \tag{2}$$

$$\lim_{m\to\infty} a_{mk} \equiv a_k \quad \text{exists} \qquad (k = 1, 2, \cdots); \tag{3}$$

$$\lim_{m\to\infty} \sum_{\nu=1}^{\infty} a_{mk_\nu} \equiv L_{\{k_\nu\}} \quad \text{exists} \qquad (\text{each } \{k_\nu\}). \tag{4}$$

We now show that this set of conditions is equivalent to the more elegant set which consists of (2) and[6] the following condition:

$$\text{There exist numbers } a_k \text{ such that } \lim_{m\to\infty} \sum_{k=1}^{\infty} | a_{mk} - a_k | = 0. \tag{5}$$

An interesting operational interpretation of this equivalence may be established as follows. We consider the Banach space of functionals W which are defined for each $x \epsilon (m)$ and for which $\| W \| = \sum_{k=1}^{\infty} | b_k |$, where $W(x) = \sum_{k=1}^{\infty} b_k s_k$ with $x = \{s_k\}$. Conditions (2), (3), and (4) now assert that $\{U_m\}$ converges *weakly* to U (B, p. 123, Theorem 2), and condition (5) asserts that $\{U_m\}$ converges *strongly* to U. Moreover, (5) shows that $U(x) = \sum_{k=1}^{\infty} a_k s_k$.

Evidently we need only to prove that the set (2), (3), (4) implies the set (2), (5). Supposing then that (2), (3), (4) are satisfied while (5) is not, we have an $\epsilon > 0$, a subsequence $\{m_\nu\}$ of $\{m\}$, and a sequence $\{M_\nu\}$ for which $M_1 < M_2 < \cdots$, such that[7]

$$\sum_{k=M(\nu)+1}^{M(\nu+1)} | a_{m_\nu k} - a_k | > 10\epsilon \quad \text{and} \quad \left(\sum_{k=1}^{M(\nu)} + \sum_{k=M(\nu+1)+1}^{\infty} \right) | a_{m_\nu k} - a_k | < \epsilon.$$

Now there are numbers $k_{\nu 1}, k_{\nu 2}, \cdots, k_{\nu\mu_\nu}$ for which[8] $M_\nu + 1 \leqq k_{\nu 1} < k_{\nu 2} < \cdots < k_{\nu\mu_\nu} \leqq M_{\nu+1}$ and

$$\left| \sum_{\mu=1}^{\mu(\nu)} (a_{m_\nu k_{\nu\mu}} - a_{k_{\nu\mu}}) \right| > 3\epsilon \qquad (\nu = 1, 3, 5, \cdots).$$

Let the sequence $\{\bar{s}_k\}$ consist of 1's in the $k_{\nu\mu}$-th places ($\mu = 1, 2, \cdots, \mu_\nu$; $\nu = 1, 3, 5, \cdots$) and 0's elsewhere. Then, for ν odd,

$$\left| \sum_{k=1}^{\infty} (a_{m_\nu k} - a_k)\bar{s}_k \right| \geqq \left| \sum_{\mu=1}^{\mu(\nu)} (a_{m_\nu k_{\mu\nu}} - a_{k_{\mu\nu}}) \right| - \left(\sum_{k=1}^{M(\nu)} + \sum_{k=M(\nu+1)+1}^{\infty} \right) | a_{m_\nu k} - a_k | > 2\epsilon;$$

and, for ν even,

$$\left| \sum_{k=1}^{\infty} (a_{m_\nu k} - a_k)\bar{s}_k \right| \leqq \left(\sum_{k=1}^{M(\nu)} + \sum_{k=M(\nu+1)+1}^{\infty} \right) | a_{m_\nu k} - a_k | < \epsilon.$$

[6] Compare H_1, p. 49, no. 21.

[7] For typographical reasons, we use notations of the form $M(\nu)$ as alternative to M_ν whenever the latter occurs as an index to a summation sign.

[8] Hamilton, *Some theorems on subsequences*, Bulletin of the American Mathematical Society, vol. 44(1938), p. 298.

Hence $\sum_{k=1}^{\infty} (a_{mk} - a_k)\bar{s}_k$ cannot converge, and thus, by (2) and (3), $\sum_{k=1}^{\infty} a_{mk}\bar{s}_k$ cannot converge. But this contradicts (4), and the proof is complete.

Passing to the general transformation, in which $l + n > 2$, and confining ourselves to the 16 types of sequences discussed in H_1, we find ourselves confronted with the problems (i) of "norming" spaces of sequences whose elements are unbounded save for subscripts sufficiently large, (ii) of making complete the spaces of ultimately regularly convergent sequences, and (iii) of discovering fundamental sets in the several spaces. (A double sequence $\{s_{\kappa\lambda}\}$ is said to be *ultimately regularly convergent* if $\lim_{\kappa\to\infty} s_{\kappa\lambda}$ and $\lim_{\lambda\to\infty} s_{\kappa\lambda}$ exist for all sufficiently large λ and κ, respectively; and an analogous definition is made for multiple sequences in general. See H_1, p. 30.) The remainder of our paper is devoted principally to the circumvention of these problems. We make henceforth free use of the definitions and symbolism of H_1 and H_2.

Circumvention of problem (i). The relevant types of sequences $\{s_k\}$ are URCRN, URCN, URC, CN, C, UB. Let V represent the generic one of these types and W the corresponding bounded type. (Thus if V is URCN, then W is BURCN.)

If $\{\sigma_m\} \epsilon$ UB, C, CN, URC, URCN, or URCRN, then (a_2) and (b_2) must be satisfied, by H_1, p. 41, nos. 2 and 4.

(For double sequences $\{s_{\kappa\lambda}\}$ which are to be transformed into sequences $\{\sigma_{\mu\nu}\}$ by means of matrices $(a_{\mu\nu\kappa\lambda})$, conditions ($a_2$), ($b_2$), and ($c_2$) become

$$(a_2)\quad \begin{aligned} a_{\mu\nu\kappa\lambda} &= 0 \quad \text{for} \quad \kappa > C_\lambda(\mu, \nu) \qquad (\mu, \nu = 1, 2, \cdots; \ \lambda = 1, 2, \cdots) \\ a_{\mu\nu\kappa\lambda} &= 0 \quad \text{for} \quad \lambda > \bar{C}_\kappa(\mu, \nu) \qquad (\mu, \nu = 1, 2, \cdots; \ \kappa = 1, 2, \cdots); \end{aligned}$$

$$(b_2)\quad \begin{aligned} a_{\mu\nu\kappa\lambda} &= 0 \quad \text{for} \quad \mu, \nu, \kappa > C_\lambda \qquad (\lambda = 1, 2, \cdots), \\ a_{\mu\nu\kappa\lambda} &= 0 \quad \text{for} \quad \mu, \nu, \lambda > \bar{C}_\kappa \qquad (\kappa = 1, 2, \cdots); \end{aligned}$$

$$(c_2)\quad \begin{aligned} a_{\mu\nu\kappa\lambda} &= 0 \quad \text{for} \quad \kappa > C_\lambda \qquad (\mu, \nu = 1, 2, \cdots; \ \lambda = 1, 2, \cdots), \\ a_{\mu\nu\kappa\lambda} &= 0 \quad \text{for} \quad \lambda > \bar{C}_\kappa \qquad (\mu, \nu = 1, 2, \cdots; \ \kappa = 1, 2, \cdots).) \end{aligned}$$

Fixing upon $x = \{s_k\} \epsilon V$, let s_k be bounded for $k > Q_x$ and take $R_x > \max C_\kappa$ $(\kappa = 1, 2, \cdots, Q_x)$. Define $y = \{t_k\}$ thus: $t_k = s_k$ for all k excepting those for which $\lambda > C_\kappa$ and $k \not> Q_x$, in which case $t_k = 0$. Then $y \epsilon W$, and

$$\sum_{k=1}^{\infty} a_{mk}\, s_k = \sum_{k=1}^{\infty} a_{mk}\, t_k \qquad \text{for } m > R_x.$$

Hence *for a matrix* (a_{mk}) *which satisfies* (a_2) *and* (b_2), *NS* $V \to UB, C, \cdots,$ *or URCRN are NS* $W \to UB, C, \cdots,$ *or URCRN, respectively.*

If $\{\sigma_m\} \epsilon$ B, BC, BCN, BURC, BURCN, RC, RCN, BURCRN, RCURN, or RCRN, (c_2) must be satisfied, by H_1, p. 42, no. 6. Fixing upon $x = \{s_k\} \epsilon V$,

let us define $y = \{t_k\}$ thus: $t_k = s_k$ for all k excepting those for which $\lambda > C_\kappa$, in which case $t_k = 0$. Then $y \epsilon W$ as before, and

$$\sum_{k=1}^{\infty} a_{mk} s_k = \sum_{k=1}^{\infty} a_{mk} t_k \qquad \text{for all } m.$$

Hence *for a matrix* (a_{mk}) *which satisfies* (c_2), *NS* $V \to B$, *BC*, $\cdots$, *or RCRN are NS* $W \to B$, *BC*, $\cdots$, *or RCRN, respectively.*

Thus problem (i) is circumvented by reducing the problem of finding NS $V \to$ UB, C, CN, URC, URCN, URCRN, B, BC, BCN, BURC, BURCN, RC, RCN, BURCRN, RCURN, or RCRN to the problem of finding NS $W \to$ UB, C, $\cdots$, or RCRN, respectively, with the provision that (a_2) and (b_2) are to be added to the latter set in the first six cases, and (c_2) in the remainder. The space (W) of sequences $\{s_k\}$ of type W is in each case normed if we take $\| x \| = \sup_k | s_k |$, where $x = \{s_k\}$. Moreover, it is complete if W is B, BC, BCN, RC, RCN, or RCRN, and the solutions of the problems of transformation in these cases are, as we shall illustrate by examples, more or less straightforward analogues of the one-dimensional problems.

If W is BURC, BURCN, BURCRN, or RCURN, the space is not complete, and we are thus led to seek another solution.

Circumvention of problem (ii). Let W represent the generic one of the above four types, and $(W(\alpha))$ the subspace of (W) whose points $\{s_k\}$ are such that $\lim_{k^1 \to \infty} s_k$ exists (and is zero for the last two types) for all $k^2 > \alpha$. Then *NS* $W \to$ *BURC, BURCN, BURCRN, or RCURN are NS* $W(\alpha) \to$ *BURC*, $\cdots$, *or RCURN for* $\alpha = 0, 1, \cdots$, *respectively.*

Thus problem (ii) is circumvented by reducing the problem of finding NS $W \to$ BURC, BURCN, BURCRN, or RCURN to the problem of finding NS $W(\alpha) \to$ BURC, $\cdots$, or RCURN for all α, respectively and, if possible, simplifying the resulting set of conditions.

Circumvention of problem (iii). For purposes of clarity, we confine ourselves henceforth to double sequences $\{s_{kl}\}$. We first introduce the following particular sequences: the sequences X_{kl} (k, $l = 1, 2, \cdots$), all of whose elements are 0 excepting that common to the k-th row and the l-th column, which is 1; the sequences X_k^r ($k = 1, 2, \cdots$), all of whose elements are 0 excepting those in the k-th row, which are 1; the sequences X_l^c ($l = 1, 2, \cdots$), all of whose elements are 0 excepting those in the l-th column, which are 1; the sequence X, all of whose elements are 1; sequences of the type Y_k^r ($k = 1, 2, \cdots$), all of whose elements are 0 excepting those in the k-th row, which are 0 and 1 in arbitrary arrangement; sequences of the type Y_l^c ($l = 1, 2, \cdots$), all of whose elements are 0 excepting those in the l-th column, which are 0 and 1 in arbitrary arrangement; and sequences of the type Y, whose elements are 0 and 1 in arbitrary arrangement.

In view of the nature of the circumventions of problems (i) and (ii) above, we see that it suffices to confine our attention to the Banach spaces (RCRN),

(RCN), (RC), (BCN), (BC), (B), (RCURN(α)), (BURCRN(α)), (BURCN(α)), and (BURC(α)). Fundamental sets for these spaces are the following. (In all cases we take $\| x \| = \sup_{k,l} | s_{kl} |$.) For (RCRN): X_{kl} $(k, l = 1, 2, \cdots)$; (RCN): X_{kl} $(k, l = 1, 2, \cdots)$, X_k^r $(k = 1, 2, \cdots)$, X_l^c $(l = 1, 2, \cdots)$; (RC): same as preceding, with X added; (BCN): X_{kl} $(k, l = 1, 2, \cdots)$, all sequences of the type Y_k^r $(k = 1, 2, \cdots)$ and of the type Y_l^c $(l = 1, 2, \cdots)$; (BC): same as preceding, with X added; (B): X_{kl} $(k, l = 1, 2, \cdots)$, all sequences of the type Y; (RCURN(α)): X_{kl} $(k, l = 1, 2, \cdots)$, X_k^r $(k = 1, 2, \cdots, \alpha)$, X_l^c $(l = 1, 2, \cdots, \alpha)$; (BURCRN($\alpha$)): X_{kl} $(k, l = 1, 2, \cdots)$, all sequences of the type Y_k^r $(k = 1, 2, \cdots, \alpha)$ and of the type Y_l^c $(l = 1, 2, \cdots, \alpha)$; (BURCN($\alpha$)): X_{kl} $(k, l = 1, 2, \cdots)$, all sequences of the type Y_k^r $(k = 1, 2, \cdots, \alpha)$ and of the type Y_l^c $(l = 1, 2, \cdots, \alpha)$, X_k^r $(k = \alpha + 1, \alpha + 2, \cdots, \alpha)$ and of the type Y_l^c $(l = 1, 2, \cdots, \alpha)$, X_k^r $(k = \alpha + 1, \alpha + 2, \cdots)$, X_l^c $(l = \alpha + 1, \alpha + 2, \cdots)$; (BURC($\alpha$)): same as preceding, with X added.

We give proofs in a few typical cases.

Proof for (*RCRN*). Setting $x_i = \sum_{k,l=1}^{i} s_{kl} X_{kl}$, we have $\lim_{i \to \infty} x_i = x$.

Proof for (*RCN*). We denote $\lim_{k \to \infty} s_{kl}$ and $\lim_{l \to \infty} s_{kl}$ by s_l^c and s_k^r, respectively. Given $\epsilon > 0$, let i be such that $| s_{kl} | < \epsilon$ for $k, l > i$, $| s_{kl} - s_l^c | < \epsilon$ for $k > i$ $(l = 1, 2, \cdots)$, and $| s_{kl} - s_k^r | < \epsilon$ for $l > i$ $(k = 1, 2, \cdots)$. Defining x_i by the equation

$$x_i = \sum_{k=1}^{i} s_k^r X_k^r + \sum_{l=1}^{i} s_l^c X_l^c + \sum_{k,l=1}^{i} (s_{kl} - s_k^r - s_l^c) X_{kl},$$

we see that $\| x_i - x \| < \epsilon$.

Proof for (*BCN*). Given $\epsilon > 0$, let i be such that $| s_{kl} | < \epsilon$ for $k, l > i$. By our work above with bounded simple sequences, we know that there exist numbers r_k^ν and c_l^μ and sequences $Y_{k\nu}^r$, $Y_{l\mu}^c$ of types Y_k^r, Y_l^c, and X_{kl} $(\nu = 1, 2, \cdots, \nu_k; \mu = 1, 2, \cdots, \mu_l; k, l = 1, 2, \cdots, i)$ such that $\| x_k^r - \sum_{\nu=1}^{\nu(k)} r_k^\nu Y_{k\nu}^r \| < \epsilon$ and $\| x_l^c - \sum_{\mu=1}^{\mu(l)} c_l^\mu Y_{l\mu}^c \| < \epsilon$ $(k, l = 1, 2, \cdots, i)$, where x_k^r and x_l^c are the simple sequences $\{x_{kl}\}$ (k fixed) and $\{x_{kl}\}$ (l fixed), respectively. Hence, defining $y_i \equiv \{t_{kl}\} = \sum_{k=1}^{i} \sum_{\nu=1}^{\nu(k)} r_k^\nu Y_{k\nu}^r + \sum_{l=1}^{i} \sum_{\mu=1}^{\mu(l)} c_l^\mu Y_{l\mu}^c$ and $x_i = y_i + \sum_{k,l=1}^{i} (s_{kl} - t_{kl}) X_{kl}$, we have $\| x_i - x \| < \epsilon$.

We continue with two illustrations of the technique used in applying operation methods to multiple sequences by deriving for double sequences $\{s_{kl}\}$ two of the sets of conditions derived by "classical" methods in H_1. We shall need the following conditions:[9]

$$(a_1) \qquad \sum_{k,l=1}^{\infty} | a_{mnkl} | < \infty \qquad (m, n = 1, 2, \cdots);$$

[9] Compare H_1, pp. 35–36.

$$\text{(b}_1\text{)} \qquad \sum_{k,l=1}^{\infty} |a_{mnkl}| < A \quad \text{for} \quad m, n > B;$$

$$\text{(c}_1\text{)} \qquad \sum_{k,l=1}^{\infty} |a_{mnkl}| < A \qquad (m, n = 1, 2, \cdots);$$

$$\text{(c}_2\text{)} \qquad \begin{aligned} a_{mnkl} &= 0 \quad \text{for} \quad k > C_l^c \qquad (m, n = 1, 2, \cdots; l = 1, 2, \cdots), \\ a_{mnkl} &= 0 \quad \text{for} \quad l > C_k^r \qquad (m, n = 1, 2, \cdots; k = 1, 2, \cdots); \end{aligned}$$

$$\text{(d}_1\text{)} \qquad \lim_{m,n\to\infty} a_{mnkl} = a_{kl} \qquad (k, l = 1, 2, \cdots);$$

$$\text{(d}_2\text{)} \qquad \begin{aligned} \lim_{m,n\to\infty} \sum_{k=1}^{\infty} a_{mnkl} &= L_l^c \qquad (l = 1, 2, \cdots), \\ \lim_{m,n\to\infty} \sum_{l=1}^{\infty} a_{mnkl} &= L_k^r \qquad (k = 1, 2, \cdots); \end{aligned}$$

$$\text{(d}_3\text{)} \qquad \lim_{m,n\to\infty} \sum_{k,l=1}^{\infty} a_{mnkl} = L;$$

(d$_4$) there exist numbers a_{kl} such that

$$\lim_{m,n\to\infty} \sum_{k=1}^{\infty} |a_{mnkl} - a_{kl}| = 0 \qquad (l = 1, 2, \cdots),$$

$$\lim_{m,n\to\infty} \sum_{l=1}^{\infty} |a_{mnkl} - a_{kl}| = 0 \qquad (k = 1, 2, \cdots);$$

$$\text{(e}_1\text{)} \qquad \begin{aligned} \lim_{m\to\infty} a_{mnkl} &= a_{nkl}^c \quad \text{for} \quad n > D \qquad (k, l = 1, 2, \cdots) \\ \lim_{n\to\infty} a_{mnkl} &= a_{mkl}^r \quad \text{for} \quad m > D \qquad (k, l = 1, 2, \cdots). \end{aligned}$$

The transformation RCRN → URC. We first show that (a$_1$) is necessary and sufficient that σ_{mn} exist for each m, n. Holding m and n fixed, we consider the additive functional U_i defined on (RCRN) by the equation $U_i(x) = \sum_{k,l=1}^{i} a_{mnkl}s_{kl}$. Since $|U_i(x)| \leqq \|x\| \sum_{k,l=1}^{i} |a_{mnkl}|$, we see by B, p. 54, Theorem 1, that U_i is a linear operation, and that $\|U_i\| \leqq \sum_{k,l=1}^{i} |a_{mnkl}|$. On the other hand, the sequence $\bar{x} = \{\bar{s}_{kl}\}$ for which $\bar{s}_{kl} = \operatorname{sgn} a_{mnkl}$ for $0 \leqq k, l \leqq i$ (sgn $0 = 1$) belongs to (RCRN) with $\|\bar{x}\| = 1$. Thus $\|U_i\| = \|U_i\| \|\bar{x}\| \geqq |U_i(\bar{x})| = \sum_{k,l=1}^{i} |a_{mnkl}|$, and we have $\|U_i\| = \sum_{k,l=1}^{i} |a_{mnkl}|$. Since $U_i(x)$ is to converge for each $x \in$ (RCRN), the necessity of (a$_1$) follows from B, p. 80, Theorem 5. Its sufficiency is obvious.

It follows that the operation U_{mn} for which $U_{mn}(x) = \sum_{k,l=1}^{\infty} a_{mnkl}s_{kl}$ is linear, and it is a triviality to see that

$$\| U_{mn} \| = \sum_{k,l=1}^{\infty} | a_{mnkl} |. \tag{6}$$

Condition (b_1) is thus seen to assert that the sequence $\{\| U_{mn} \|\}$ is ultimately bounded, a circumstance which follows from B, p. 80, Theorem 5, in view of the fact that $\lim_{i\to\infty} \sigma_{m_i n_i}$ must exist for each pair of sequences $\{m_i\}$ and $\{n_i\}$ for which $\lim_{i\to\infty} m_i = \lim_{i\to\infty} n_i = \infty$.

The necessity of (d_1) follows from the facts that $X_{kl} \in$ (RCRN) and that $U_{mn}(X_{kl}) = a_{mnkl}$.

The conditions (a_1), (b_1), and (d_1) insure the convergence of $\sigma_{mn} = U_{mn}(x)$ for each x, in view of B, p. 79, Theorem 3, since, by (b_1), the sequence $\{\| U_{m_i n_i} \|\}$ is bounded for each pair of sequences $\{m_i\}$ and $\{n_i\}$ as characterized above.

Suppose next that (e_1)—and, for definiteness, the first part—is not satisfied. Then there exist sequences $\{k_i\}$, $\{l_i\}$, $\{n_i\}$, and $\{m_{ij}\}$ such that $\lim_{i\to\infty} n_i = \infty$, $\lim_{j\to\infty} m_{ij} = \infty$ for $i = 1, 2, \cdots$, and $\lim_{j\to\infty} a_{m_{ij}n_ik_il_i}$ fails to exist for $i = 1, 2, \cdots$. Defining $U_{ij} = U_{m_{ij}n_i}$, we thus see that $\lim_{j\to\infty} U_{ij}(X_{k_il_i})$ fails to exist for $i = 1, 2, \cdots$. Hence, by B, p. 24, Theorem 6, there exists a sequence $\{\bar{s}_{kl}\} \equiv \bar{x} \in$ (RCRN) for which $\lim_{j\to\infty} U_{ij}(\bar{x})$ fails to exist for $i = 1, 2, \cdots$. Thus $\sigma_{mn} \equiv \sum_{k,l=1}^{\infty} a_{mnkl}\bar{s}_{kl}$ is not urc, contrary to our requirements.

The conditions (a_1), (b_1), and (e_1) insure the existence of $\lim_{m\to\infty} \sigma_{mn}$ for $n > B, D$ and of $\lim_{n\to\infty} \sigma_{mn}$ for $m > B, D$, by B, p. 79, Theorem 3.

Finally, then,[10] NS RCRN $\to$ URC are (a_1), (b_1), (d_1), and (e_1).

The transformation URC → C. As shown before, we need merely derive conditions necessary and sufficient that BURC(α) $\to$ C for each α, reduce these conditions if possible, and add (a_2) and (b_2). Defining U_{mn} as above, we find, as there, that $\| U_{mn} \| = \sum_{k,l=1}^{\infty} | a_{mnkl} |$ and that (a_1), (b_1), and (d_1) must be satisfied.

Similarly, we establish the necessity of (d_3);

$$(d_{2\alpha}) \qquad \begin{aligned} &\lim_{m,n\to\infty} \sum_{k=1}^{\infty} a_{mnkl} = L_l^c \qquad (l = \alpha + 1, \alpha + 2, \cdots), \\ &\lim_{m,n\to\infty} \sum_{l=1}^{\infty} a_{mnkl} = L_k^r \qquad (k = \alpha + 1, \alpha + 2, \cdots); \end{aligned}$$

[10] Compare H_1, p. 50, no. 43, and p. 58, §7.

and

$$(7)\qquad \begin{aligned} &\lim_{m,n\to\infty}\sum_{\nu=1}^{\infty} a_{mnk_\nu l} = L_l^c(\{k_\nu\}) &&(l = 1, 2, \cdots, \alpha;\text{ each } \{k_\nu\}),\\ &\lim_{m,n\to\infty}\sum_{\nu=1}^{\infty} a_{mnkl_\nu} = L_k^r(\{l_\nu\}) &&(k = 1, 2, \cdots, \alpha;\text{ each } \{l_\nu\}).\end{aligned}$$

The sufficiency of the conditions (a_1), (b_1), (d_1), ($d_{2\alpha}$), (d_3), and (7) follows as before. Moreover, we find, somewhat as in the case of the simple sequence transformation treated above, that we may replace (7) in this set of necessary and sufficient conditions by

$$(d_{4\alpha})\qquad \begin{aligned} &\lim_{m,n\to\infty}\sum_{k=1}^{\infty} | a_{mnkl} - a_{kl} | = 0 &&(l = 1, 2, \cdots, \alpha),\\ &\lim_{m,n\to\infty}\sum_{l=1}^{\infty} | a_{mnkl} - a_{kl} | = 0 &&(k = 1, 2, \cdots, \alpha).\end{aligned}$$

But the holding of ($d_{2\alpha}$) and ($d_{4\alpha}$) for $\alpha = 0, 1, \cdots$ is equivalent to the holding of (d_2) and (d_4). And since (d_4) and (b_1) imply (d_2) and (d_1), it follows that *NS BURC → C are* (a_1), (b_1), (d_3), *and* (d_4). Finally, then,[11] *NS URC → C are* (a_1), (a_2), (b_1), (b_2), (d_3), *and* (d_4).

Concluding, we mention that in certain transformations involving the existence of the limit (or limits) of σ_{mn} it is a simple matter to evaluate this (or these) limit(s) by means of operational methods. Thus in the former of the two transformations considered above, we may proceed as follows to find the value of $\sigma \equiv \lim_{m,n\to\infty} \sigma_{mn}$.

Defining $U(x) \equiv \lim_{m,n\to\infty} U_{mn}(x)$, we have $\sigma = U(x) = \lim_{i\to\infty} U_{ii}(x)$, whence, by (6) and ($b_1$), we see easily that U is linear. If, then, we set $x_j = \sum_{k,l=1}^{j} s_{kl}X_{kl}$, we have $x = \lim_{j\to\infty} x_j$ so that $\sigma = \lim_{j\to\infty} U(x_j)$. But, by ($d_1$), we have $U(x_j) = \sum_{k,l=1}^{j} a_{kl}s_{kl}$. Hence, finally, $\sigma = \sum_{k,l=1}^{\infty} a_{kl}s_{kl}$.

Michigan State College and Pomona College.

[11] Compare H_1, p. 49, no. 14; p. 58, §7, and p. 38, (.12).

FUNCTIONS WITH POSITIVE DERIVATIVES

By R. P. Boas, Jr.

D. V. Widder has called my attention to the fact that a function having all its derivatives of even order non-negative in an interval is necessarily analytic there. This is a special case of results which have been stated, without detailed proof, by S. Bernstein;[1] it is useful in the theory of Laplace integrals.[2] In the first part of this note, I give an elementary proof of the theorem, and a proof, by methods differing from Bernstein's, of Bernstein's general theorem. This is

Theorem 1. *Let* $\{n_p\}$ $(p = 1, 2, \cdots)$ *be an increasing infinite sequence of positive integers such that* n_{p+1}/n_p *is bounded. If* $f(x)$ *is of class*[3] C^∞ *in* $a < x < b$, *and has the property that for each* p $(p = 1, 2, \cdots)$ $f^{(n_p)}(x)$ *does not change sign in* $a < x < b$, *then* $f(x)$ *is analytic in* $a < x < b$.

The greater part of this note is devoted to showing that Theorem 1 is, in a certain direction, the best possible result. I construct, for any sequence $\{n_p\}$ such that $n_{p+1}/n_p \to \infty$, a function whose n_p-th derivatives are positive in an interval, but which is not analytic in the interval. (The case where lim sup $(n_{p+1}/n_p) = \infty$, lim inf $(n_{p+1}/n_p) < \infty$ is left open.) More precisely, I prove

Theorem 2. *Let* $\{n_p\}$ $(p = 1, 2, \cdots)$ *be an increasing infinite sequence of positive integers such that* $\lim_{p\to\infty} n_{p+1}/n_p = \infty$. *Then there is a function* $f(x)$, *of class* C^∞ *in* $-1 < x < 1$, *such that* $f(x) > 0$ *in* $-1 < x < 1$, *and*

$$f^{(n_p)}(x) > 0 \quad (-1 < x < 1; p = 1, 2, \cdots), \tag{1}$$

$$f(x) \textit{ is not analytic in } -1 < x < 1. \tag{2}$$

The function $f(x)$ will be defined by means of its development in a series of Chebysheff polynomials.

In Theorem 1, it could equally well be supposed that $f(x)$, instead of having n_p-th derivatives which do not change sign, is continuous and has n_p-th dif-

Received September 21, 1940; presented to the American Mathematical Society, September 10, 1940.

[1] S. Bernstein, *Leçons sur les Propriétés Extrémales et la Meilleure Approximation des Fonctions Analytiques d'une Variable Réelle*, Paris, 1926, pp. 196–197.

[2] D. V. Widder, *Necessary and sufficient conditions for the representation of a function by a doubly infinite Laplace integral*, Bulletin of the American Mathematical Society, vol. 40(1934), pp. 321–326.

[3] A function is of class C^n $(n = 1, 2, \cdots)$ if it has a continuous n-th derivative; of class C^∞ if of class C^n for every n.

ferences which do not change sign. In fact, the latter hypothesis implies the former, by a theorem of T. Popoviciu.[4]

1. **Proof of Theorem 1.** We require Lemmas 1 and 2, Lemma 3 for $n_p = 2p$, and Lemma 4 for general n_p. Lemmas 1 and 2 are slight modifications of lemmas of E. Landau.[5]

LEMMA 1. *In* $0 \leqq x \leqq h$ *let* $f(x) \in C^n$ $(n \geqq 2)$, $|f(x)| \leqq 1, f^{(n)}(x) \geqq 0$. *Then*

$$f^{(n-1)}(0) \leqq \left(\frac{2}{h}\right)^{n-1} (n-1)^{n-1}.$$

Let $\delta = h/(n-1)$. Then

$$\Delta_\delta^{n-1} f(0) = \int_0^\delta dx_1 \int_{x_1}^{x_1+\delta} dx_2 \cdots \int_{x_{n-2}}^{x_{n-2}+\delta} f^{(n-1)}(x_{n-1})\, dx_{n-1}. \tag{1.1}$$

In the integral on the right, $f^{(n-1)}(x_{n-1}) \geqq f^{(n-1)}(0)$ (because $f^{(n)}(x) \geqq 0$), so that the right side of (1.1) is not less than $\delta^{n-1} f^{(n-1)}(0)$. The left side of (1.1) is at most

$$\sum_{i=0}^{n-1} \binom{n-1}{i} = 2^{n-1}.$$

Hence

$$f^{(n-1)}(0) \leqq \left(\frac{2}{\delta}\right)^{n-1} = \left(\frac{2}{h}\right)^{n-1} (n-1)^{n-1}.$$

LEMMA 2. *In* $-1 \leqq x \leqq 1$, *let* $f(x) \in C^n$ $(n \geqq 2)$, $|f(x)| \leqq 1, f^{(n)}(x) \geqq 0$. *Then in* $|x| \leqq 1 - \delta$ $(0 < \delta < 1)$

$$|f^{(n-1)}(x)| \leqq \left(\frac{2}{\delta}\right)^{n-1} (n-1)^{n-1}. \tag{1.2}$$

If $-1 \leqq \xi \leqq 1 - \delta$, Lemma 1 can be applied to $f(x + \xi)$, as a function of x, in $0 \leqq x \leqq \delta$. The conclusion is

$$f^{(n-1)}(\xi) \leqq \left(\frac{2}{\delta}\right)^{n-1} (n-1)^{n-1}. \tag{1.3}$$

If $-1 + \delta < \xi \leqq 1$, Lemma 1 can be applied to $(-1)^n f(-x + \xi)$, as a function of x, in $0 \leqq x \leqq \delta$. The conclusion is

$$f^{(n-1)}(\xi) \geqq \left(\frac{2}{\delta}\right)^{n-1} (n-1)^{n-1}. \tag{1.4}$$

Inequalities (1.3) and (1.4) imply (1.2).

[4] T. Popoviciu, *Sur l'approximation des fonctions convexes d'ordre supérieur*, Mathematica (Cluj), vol. 8(1934), pp. 1–85; pp. 54–58. Another proof has been given by Boas and Widder, *Functions with positive differences*, this Journal, vol. 7(1940), pp. 496–503.

[5] E. Landau, *Über einen Satz von Herrn Esclangon*, Mathematische Annalen, vol. 102(1929–30), pp. 177–188; 184–185.

LEMMA 3.[6] *In* $-h \leqq x \leqq h$, *let* $f(x) \in C^2$, $|f(x)| \leqq A$, $|f''(x)| \leqq B$, $B/A > 4/h^2$. *Then* $|f'(x)| \leqq 2(AB)^{\frac{1}{2}}$ *in* $-h \leqq x \leqq h$.

If $|\delta| \leqq h$ and x and $x + \delta$ are both in the closed interval $(-h, h)$, Taylor's theorem with remainder of order 2 yields

$$f'(x) = \frac{1}{\delta}[f(x+\delta) - f(x)] - \frac{1}{2}\delta f''(x+\theta\delta), \qquad 0 < \theta < 1;$$

$$|f'(x)| \leqq \frac{2A}{\delta} + \frac{\delta B}{2}.$$

If $B/A \geqq 4/h^2$, $2(A/B)^{\frac{1}{2}} \leqq h$, and we may take $\delta = -2(A/B)^{\frac{1}{2}}\operatorname{sgn} x$. Then

$$|f'(x)| \leqq 2(AB)^{\frac{1}{2}}.$$

LEMMA 4. *In* $-h \leqq x \leqq h$, *let* $f(x) \in C^n$ $(n \geqq 2)$, $|f(x)| \leqq M_0$, $|f^{(n)}(x)| \leqq M_n$. *Then in* $-h \leqq x \leqq h$

$$|f^{(k)}(x)| \leqq 4e^{2k}\left(\frac{n}{k}\right)^k M_0^{1-k/n} M_n'^{k/n} \quad (k = 1, 2, \cdots, n-1),$$

where M_n' *denotes the larger of* M_n *and* $M_0 n! h^{-n}$.

Lemma 4 was proved by A. Gorny; H. Cartan has given a similar result.[7]

Proof of Theorem 1 *for* $n_p = 2p$. Without loss of generality we assume that $f(x)$ is of class C^∞ in $-1 \leqq x \leqq 1$, and that $|f(x)| \leqq 1$. By Lemma 2,

$$(1.5) \qquad |f^{(2p-1)}(x)| \leqq \left(\frac{2}{\delta}\right)^{2p-1}(2p-1)^{2p-1} \qquad (|x| \leqq 1 - \delta;\ p = 1, 2, \cdots).$$

By Lemma 3, applied to $f^{(2p-1)}(x)$ and $f^{(2p+1)}(x)$,

$$|f^{(2p)}(x)| \leqq 2\left(\frac{2}{\delta}\right)^{2p}\{(2p-1)^{2p-1}(2p+1)^{2p+1}\}^{\frac{1}{2}},$$

provided that

$$\left(\frac{2}{\delta}\right)^2 \frac{(2p+1)^{2p+1}}{(2p-1)^{2p-1}} \geqq \frac{4}{(1-\delta)^2};$$

this will be true if $2p - 1 \geqq \delta/(1-\delta)$, and so for sufficiently large p. We thus have, for large p,

$$|f^{(2p)}(x)| \leqq 2\left(\frac{2}{\delta}\right)^{2p}(2p+1)(2p)^{2p-1},$$

[6] This is a variant of a known lemma. Cf. E. Landau, *Einige Ungleichungen für zweimal differentiierbare Funktionen*, Proceedings of the London Mathematical Society, (2), vol. 13(1913), pp. 43–49; T. Carleman, *Les Fonctions Quasi Analytiques*, Paris, 1926, p. 12.

[7] A. Gorny, *Contribution à l'étude des fonctions dérivables d'une variable réelle*, Acta Mathematica, vol. 71(1939), pp. 317–358.

Cartan's results are stated by H. Cartan and S. Mandelbrojt, *Solution du problème d'équivalence des classes de fonctions indéfiniment dérivables*, Acta Mathematica, vol. 72(1940), pp. 31–49; p. 33.

since $\{(2p-1)(2p+1)\}^{\frac{1}{2}} < 2p$; or

$$|f^{(2p)}(x)| \leqq 4\left(\frac{4p}{\delta}\right)^{2p} \qquad (|x| \leqq 1-\delta). \tag{1.6}$$

From (1.5) and (1.6)

$$|f^{(k)}(x)| \leqq 4\left(\frac{2k}{\delta}\right)^{k} \qquad (|x| \leqq 1-\delta)$$

for all sufficiently large integers k. By Stirling's formula, it follows that

$$\limsup_{k\to\infty}\left|\frac{f^{(k)}(x)}{k!}\right|^{1/k} \leqq \frac{2e}{\delta},$$

uniformly in $|x| \leqq 1-\delta$. Hence $f(x)$ is analytic in $(-1, 1)$, and the radius of convergence of its Taylor series about the point x is at least $1/(2e)$ times the distance from x to the nearer endpoint of $(-1, 1)$.

Proof of Theorem 1 *for general* n_p. Without loss of generality we assume that $f(x)$ is of class C^∞ in $-1 \leqq x \leqq 1$, and that $|f(x)| \leqq 1$. By Lemma 2,

$$|f^{(n_p-1)}(x)| \leqq \left(\frac{2}{\delta}\right)^{n_p-1}(n_p-1)^{n_p-1} \qquad (|x| \leqq 1-\delta;\, p = 1, 2, \cdots).$$

By Lemma 4, for $|x| \leqq 1-\delta$ and $k = n_{p-1}, n_{p-1}+1, \cdots, n_p-2$, either

$$|f^{(k)}(x)| \leqq 4\left(\frac{2e^2}{k\delta}\right)^{k}(n_p-1)^{2k}$$

or

$$\begin{aligned}|f^{(k)}(x)| &\leqq 4\left(\frac{e^2}{k\delta}\right)^{k}(n_p-1)^{k}\{(n_p-1)!\}^{k/(n_p-1)}\\ &\leqq 4\left(\frac{e}{k\delta}\right)^{k}(n_p-1)^{2k}\{2\pi e^{\frac{1}{2}}(n_p-1)\}^{\frac{1}{2}k/(n_p-1)}.\end{aligned}$$

In either case, it follows that

$$\limsup_{k\to\infty}\left|\frac{f^{(k)}(x)}{k!}\right|^{1/k} \leqq \frac{2e^3}{\delta}\limsup_{p\to\infty}\left(\frac{n_p}{n_{p-1}}\right)^2,$$

uniformly in $|x| \leqq 1-\delta$. Hence $f(x)$ is analytic in $(-1, 1)$; and, if lim sup $(n_{p+1}/n_p) = c$, the radius of convergence of the Taylor series of $f(x)$ about any point x is at least[8] $1/(2e^3c^2)$ times the distance from x to the nearer endpoint of $(-1, 1)$.

2. **Proof of Theorem 2.** The function $f(x)$ will be defined by a series of Chebysheff polynomials (the $T_n(x)$ of the following definition).

[8] This value can be improved by use of the estimates, also given by Gorny and Cartan, for the magnitude of $|f^{(k)}(0)|$ under the hypotheses of Lemma 4.

DEFINITION. For $n = 0, 1, 2, \cdots$,

$$T_n(x) = \cos(n \cos^{-1} x) = \tfrac{1}{2}\{[x + \sqrt{(x^2 - 1)}]^n + [x - \sqrt{(x^2 - 1)}]^n\}.$$

For $n = 1, 2, \cdots$,

$$S_n(x) = \tfrac{1}{2}[T_n(x) + T_{n-1}(x)], \qquad S_0(x) = 1. \tag{2.1}$$

LEMMA 5.[9] *$T_n(x)$ is a polynomial of degree n, in which the coefficient of x^n is 2^{n-1} $(n = 1, 2, \cdots)$. When n is even, $|T_n(0)| = 1$ and $T_n'(0) = 0$; when n is odd, $T_n(0) = 0$ and $|T_n'(0)| = n$.*

Lemmas 6, 7, and 8 are well known,[10] but since I cannot quote explicit proofs, I outline proofs here.

LEMMA 6. *For $k = 1, 2, \cdots, n$,*

$$|S^{(k)}(0)| \leqq n^k \qquad (n = 1, 2, \cdots). \tag{2.2}$$

$T_n(x)$ satisfies the differential equations[11]

$$(x^2 - 1)T_n^{(k+1)}(x) + (2k - 1)xT_n^{(k)}(x) = \{n^2 - (k - 1)^2\}T_n^{(k-1)}(x) \qquad (k = 1, 2, \cdots). \tag{2.3}$$

For $x = 0$, we obtain by successive applications of (2.3)

$$\begin{aligned} T_n^{(k+1)}(0) &= \{n^2 - (k-1)^2\}T_n^{(k-1)}(0) = \cdots \\ &= \begin{cases} \{n^2 - (k-1)^2\}\{n^2 - (k-3)^2\} \cdots n^2 T_n(0) & (k \text{ odd}) \\ \{n^2 - (k-1)^2\}\{n^2 - (k-3)^2\} \cdots (n^2 - 1)T_n'(0) & (k \text{ even}). \end{cases} \end{aligned} \tag{2.4}$$

From (2.4), Lemma 5, and (2.1), the conclusion follows at once.

LEMMA 7.[12] *For $1 \leqq k \leqq \frac{1}{2}n$,*

$$|S_n^{(k)}(0)| \geqq (c_3 n)^k. \tag{2.5}$$

From (2.4) and Lemma 5 we have, if n is even, k is odd, and $k \leqq \frac{1}{2}n$,

$$|T_n^{(k+1)}(0)| \geqq \{n^2 - (k-1)^2\}^{\frac{1}{2}(k+1)} \geqq (c_1 n)^{k+1};$$

if n is odd and k is even,

$$|T_n^{(k+1)}(0)| \geqq n\{n^2 - (k-1)^2\}^{\frac{1}{2}k} \geqq (c_2 n)^{k+1};$$

if n and k have the same parity, $T_n^{(k+1)}(0) = 0$. Relation (2.5) follows from these inequalities.

[9] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, Berlin, 1925, vol. 2, p. 75.

[10] Cf. Cartan and Mandelbrojt, op. cit., p. 44.

[11] W. Markoff, *Über Polynome, die in einem gegebenen Intervalle möglichst wenig von Null abweichen*, Mathematische Annalen, vol. 77(1916), pp. 213–258; p. 231.

[12] $c_1, c_2, \cdots, c_7$ denote positive constants.

LEMMA 8. *For* $k = 1, 2, \cdots, n$, *and* $-1 \leqq x \leqq 1$,

$$|S_n^{(k)}(x)| \leqq \left(\frac{c_5 n^2}{k}\right)^k. \tag{2.6}$$

It was shown by W. Markoff[13] that $|T_n^{(k)}(x)|$ attains its maximum in $(-1, 1)$ at $x = 1$, where it has the value

$$\frac{n^2(n^2-1^2)(n^2-2^2)\cdots(n^2-(k-1)^2)}{1\cdot 3\cdot 5\cdots(2k-1)} \leqq \frac{n^{2k}k!2^k}{(2k)!}$$
$$\leqq c_4\left(\frac{n^2 e}{2k}\right)^k,$$

by Stirling's formula; (2.6) follows.

We now show that for suitably chosen numbers ϵ and L $(0 < \epsilon < \frac{1}{2}, L \geqq 0)$ the function

$$f(x) = L + \sum_{p=1}^{\infty} \lambda_p S_{n_p}(x), \tag{2.7}$$

with

$$\lambda_1 = 1, \qquad \frac{\lambda_{p+1}}{\lambda_p} = \left(\frac{n_p}{n_{p+1}}\right)^{2n_p} \epsilon^{n_p} \qquad (p = 1, 2, \cdots)$$

has the properties asserted in Theorem 2.

LEMMA 9. *For* $r = 1, 2, \cdots$,

$$\sum_{p=r+1}^{\infty} \lambda_p n_p^{2n_r} \leqq 2\lambda_r n_r^{2n_r} \epsilon^{n_r}. \tag{2.8}$$

Since $0 < \epsilon < 1$ and $0 < n_p/n_{p+1} < 1$, we have for $p = 2, 3, \cdots$

$$\frac{\lambda_p}{\lambda_{p-1}} = \left(\frac{n_{p-1}}{n_p}\right)^{2n_r} \epsilon^{n_r} \qquad (1 \leqq r \leqq p-1),$$

$$\lambda_p n_p^{2n_r} \leqq \lambda_{p-1} n_{p-1}^{2n_r} \epsilon^{n_r} \qquad (1 \leqq r \leqq p-1).$$

By repeated applications of this inequality we obtain

$$\lambda_p n_p^{2n_r} \leqq \lambda_r n_r^{2n_r} \epsilon^{n_r(p-r)} \qquad (1 \leqq r \leqq p-1).$$

Consequently

$$\begin{aligned}\sum_{p=r+1}^{\infty} \lambda_p n_p^{2n_r} &\leqq \lambda_r n_r^{2n_r} \sum_{p=r+1}^{\infty} \epsilon^{n_r(p-r)} \\ &= \frac{\lambda_r n_r^{2n_r} \epsilon^{n_r}}{1-\epsilon^{n_r}} \\ &\leqq \frac{\lambda_r n_r^{2n_r} \epsilon^{n_r}}{1-(\frac{1}{2})^{n_r}} \\ &\leqq 2\epsilon^{n_r} \lambda_r n_r^{2n_r} \qquad (r = 1, 2, \cdots).\end{aligned}$$

[13] W. Markoff, op. cit., p. 250.

We can now establish statement (1) of Theorem 2.

Because of (2.8), $\sum \lambda_p$ converges, so that (since $|S_n(x)| \leqq 1$) $f(x)$ is well-defined. The formal series for $f^{(k)}(x)$ is

$$\sum_{p=1}^{\infty} \lambda_p S_{n_p}^{(k)}(x),$$

which, by Lemma 8, is dominated by

$$\sum_{n_p \geqq k} \lambda_p \left(\frac{c_5 n_p^2}{p}\right)^k;$$

this dominant series converges by Lemma 9. Hence $f(x)$ has derivatives of all orders, obtained by formal differentiation of (2.7). We have

$$f^{(n_r)}(x) = S_{n_r}^{(n_r)}(x) + \sum_{p=r+1}^{\infty} \lambda_p S_{n_p}^{(n_r)}(x) \qquad (r = 1, 2, \cdots).$$

Now

$$S_n^{(n)}(x) = \tfrac{1}{2}[T_n^{(n)}(x) + T_{n-1}^{(n)}(x)] = 2^{n-2} n! \qquad (n = 1, 2, \cdots),$$

so that

$$f^{(n_r)}(x) = \lambda_r n_r! 2^{n_r - 2} + \sum_{p=r+1}^{\infty} \lambda_p S_{n_p}^{(n_r)}(x) \qquad (r = 1, 2, \cdots). \tag{2.9}$$

We show now that ϵ can be chosen so small that

$$\left| \sum_{p=r+1}^{\infty} \lambda_p S_{n_p}^{(n_r)}(x) \right| \leqq \tfrac{1}{2}\lambda_r n_r! 2^{n_r - 2} \qquad (-1 \leqq x \leqq 1; r = 1, 2, \cdots). \tag{2.10}$$

By Lemmas 8 and 9 we have

$$\begin{aligned} \left| \sum_{p=r+1}^{\infty} \lambda_p S_{n_p}^{(n_r)}(x) \right| &\leqq \sum_{p=r+1}^{\infty} \lambda_p c_5^{n_r} \left(\frac{n_p^2}{n_r}\right)^{n_r} \\ &\leqq 2 c_5^{n_r} n_r^{-n_r} \lambda_r n_r^{2n_r} \epsilon^{n_r} \\ &\leqq c_6^{n_r} \epsilon^{n_r} n_r! \lambda_r \end{aligned}$$

(by Stirling's formula). To make (2.10) true, we have only to choose ϵ less than $1/(4c_6)$. We choose ϵ so small that also

$$\epsilon < \tfrac{1}{4} c_3, \tag{2.11}$$

where c_3 is the constant of Lemma 7.

By (2.9) and (2.10), $f^{(n_r)}(x) > 0$ $(-1 \leqq x \leqq 1; r = 1, 2, \cdots)$, so that (1) is established. By choosing L sufficiently large, we can make $f(x) > 0$ $(-1 \leqq x \leqq 1)$.

To establish (2) it is enough to show that

$$\limsup_{n \to \infty} \left| \frac{f^{(n)}(0)}{n!} \right|^{1/n} = \infty, \tag{2.12}$$

so that $f(x)$ is not analytic at $x = 0$. We denote by m_p the integral part of $(n_p n_{p+1})^{\frac{1}{2}}$; we shall show that

$$\lim_{p\to\infty} \left| \frac{f^{(m_p)}(0)}{m_p!} \right|^{1/m_p} = \infty; \tag{2.13}$$

this will establish (2.12).

There is a positive integer q such that if $p > q$, $n_{p+1}/n_p > 4$. For $p > q$ we then have

$$n_p^{\frac{1}{2}} < \tfrac{1}{2} n_{p+1}^{\frac{1}{2}},$$

$$m_p \leqq (n_p n_{p+1})^{\frac{1}{2}} < \tfrac{1}{2} n_{p+1};$$

and

$$n_p < (n_p n_{p+1})^{\frac{1}{2}} - 1 \leqq m_p;$$

so that

$$n_p < m_p < \tfrac{1}{2} n_{p+1}. \tag{2.14}$$

We have

$$f^{(m_p)}(0) = \lambda_{p+1} S_{n_{p+1}}^{(m_p)}(0) + \sum_{k=p+2}^{\infty} \lambda_k S_{n_k}^{(m_p)}(0);$$

and, by Lemmas 6 and 7,

$$f^{(m_p)}(0) = \lambda_{p+1}(c_3 n_{p+1})^{m_p} - \sum_{k=p+2}^{\infty} \lambda_k n_k^{m_p}. \tag{2.15}$$

Now

$$\frac{\lambda_{p+1}}{\lambda_p} = \left(\frac{n_p}{n_{p+1}}\right)^{2n_p} \epsilon^{n_p} \leqq \left(\frac{n_p}{n_{p+1}}\right)^{n_p} \epsilon^{n_p};$$

$$\frac{\lambda_{p+1}}{\lambda_p} \leqq \left(\frac{n_p}{n_{p+1}}\right)^{m_r} \epsilon^{m_r} \qquad (p - 1 \geqq r > q),$$

$$\lambda_{p+1} n_{p+1}^{m_r} \leqq \epsilon^{m_r} \lambda_p n_p^{m_r};$$

hence

$$\lambda_k n_k^{m_p} \leqq \epsilon^{(k-p-1)m_p} \lambda_{p+1} n_{p+1}^{m_p} \qquad (k \geqq p + 1 > q + 1),$$

and

$$\begin{aligned} \sum_{k=p+2}^{\infty} \lambda_k n_k^{m_p} &\leqq \lambda_{p+1} n_{p+1}^{m_p} \frac{\epsilon^{m_p}}{1 - \epsilon^{m_p}} \\ &\leqq 2\epsilon^{m_p} \lambda_{p+1} n_{p+1}^{m_p} \\ &\leqq \tfrac{1}{2} \lambda_{p+1} (c_3 n_{p+1})^{m_p} \end{aligned}$$

since $0 < \epsilon < \frac{1}{2}$, and $\epsilon < \frac{1}{4} c_3$ by (2.11).

From (2.15) we now obtain

$$|f^{(m_p)}(0)| \geqq \tfrac{1}{2}\lambda_{p+1}(c_3 n_{p+1})^{m_p},$$

$$\left|\frac{f^{(m_p)}(0)}{m_p!}\right|^{1/m_p} \geqq c_7 \lambda_{p+1}^{1/m_p} \frac{n_{p+1}}{m_p}.$$

Since

$$\frac{n_{p+1}}{m_p} \geqq \frac{n_{p+1}}{(n_p n_{p+1})^{\frac{1}{2}} - 1} \to \infty \qquad (p \to \infty),$$

(2.13) will follow if we show that

$$\lambda_{p+1}^{1/m_p} \to 1 \qquad (p \to \infty). \tag{2.16}$$

Since $\lambda_1 = 1$ and

$$\frac{\lambda_{p+1}}{\lambda_p} = \left(\frac{n_p}{n_{p+1}}\right)^{2n_p} \epsilon^{n_p},$$

it follows that

$$\begin{aligned} \lambda_{p+1} &= \left(\frac{n_p}{n_{p+1}}\right)^{2n_p} \left(\frac{n_{p-1}}{n_p}\right)^{2n_{p-1}} \cdots \left(\frac{n_1}{n_2}\right)^{2n_1} \epsilon^{n_p + n_{p-1} + \cdots + n_1} \\ &= \left(\frac{n_p}{n_{p+1}}\right)^{2n_p} P^2 \epsilon^{n_p + n_{p-1} + \cdots + n_1}, \end{aligned}$$

with

$$P = \prod_{r=1}^{p-1} \left(\frac{n_r}{n_{r+1}}\right)^{n_r}.$$

Now

$$\log \left(\frac{n_p}{n_{p+1}}\right)^{2n_p/m_p} = 2\,\frac{n_p}{m_p} \log \frac{n_p}{n_{p+1}} \to 0 \qquad (p \to \infty),$$

so that

$$\lim_{p\to\infty} \left(\frac{n_p}{n_{p+1}}\right)^{2n_p/m_p} = 1. \tag{2.17}$$

For $p > q + 1$, $n_p/n_{p-1} > 4$, and consequently

$$n_p > 2/n_{p-1} > \cdots > 4^{p-q-1} n_{q+1},$$

$$\begin{aligned} n_p + n_{p-1} + \cdots + n_1 &\leqq n_p \left(1 + \tfrac{1}{4} + \cdots + \frac{1}{4^{p-q-1}}\right) + q n_{q+1} \\ &= \tfrac{4}{3} n_p (1 - 4^{q-p}) + q n_{q+1}; \end{aligned}$$

$$\begin{aligned} \frac{n_p + n_{p-1} + \cdots + n_1}{m_p} &\leqq \frac{4n_p}{3m_p} + \frac{q n_{q+1}}{m_p} \\ &\to 0 \qquad (p \to \infty). \end{aligned}$$

Hence

$$\lim_{p\to\infty} \epsilon^{(n_p+n_{p-1}+\cdots+n_1)/m_p} = 1. \tag{2.18}$$

Finally, we consider the expression P. We have

$$0 \leqq -\log P = \sum_{r=1}^{p-1} n_r (\log n_{r+1} - \log n_r)$$

$$= \sum_{r=1}^{p-1} e^{\sigma_r}(\sigma_{r+1} - \sigma_r),$$

where we write $\sigma_r = \log n_r$. Now

$$\sum_{r=1}^{p-1} e^{\sigma_r}(\sigma_{r+1} - \sigma_r) \leqq \int_{\sigma_1}^{\sigma_p} e^x\, dx < e^{\sigma_p} = n_p,$$

and hence

$$0 \leqq -\frac{\log P}{m_p} < \frac{n_p}{m_p} \to 0 \qquad (p \to \infty).$$

Consequently

$$\lim_{p\to\infty} P^{1/m_p} = 1. \tag{2.19}$$

From (2.17), (2.18), and (2.19), (2.16) follows. This completes the proof.[14]

DUKE UNIVERSITY.

[14] **(Added in proof.)** I. J. Schoenberg has proved that if $f^{(2n)}(x) \geqq 0$ $(n = 0, 1, 2, \ldots)$ in (a, b) the Taylor series of $f(x)$ about the point x converges out to the nearer endpoint of (a, b). This gives, for this special case, more than is established in Theorem 1.

SIMULTANEOUS REPRESENTATION IN A QUADRATIC AND LINEAR FORM

By Gordon Pall

1. **Reduction to a single equation.** Let $c_1, \cdots, c_s, a, b$ be given integers. Consider the solvability in integers x_i of the pair of equations

$$(1) \qquad c_1x_1^2 + \cdots + c_sx_s^2 = a, \qquad c_1x_1 + \cdots + c_sx_s = b.$$

Set $u = c_1 \cdots c_s$, $t = c_1 + \cdots + c_s$, and assume $tu \neq 0$. The identity

$$(2) \qquad \left(\sum c_i\right)\left(\sum c_ix_i^2\right) - \left(\sum c_ix_i\right)^2 = \sum_{i<k}^{1,\cdots,s} c_ic_k(x_i - x_k)^2$$

suggests introducing the new variables

$$(3) \qquad y_j = x_1 - x_j \qquad (j = 2, \cdots, s),$$

whence $x_i - x_k = y_k - y_i$. Then by (1) and (2),

$$(4) \qquad ta - b^2 = \phi(y_2, \cdots, y_s),$$

where ϕ is the quadratic form, in $s - 1$ variables,

$$(5) \qquad \sum_{j}^{2,\cdots,s} c_j(t - c_j)y_j^2 - 2\sum_{j<k}^{2,\cdots,s} c_jc_ky_jy_k.$$

2. The author[1] treated a more general pair of equations $a = q(x_1, \cdots, x_s)$, $b = l(x_1, \cdots, x_s)$ in 1931, the coefficients of q and l being unrelated. The present article was suggested by recent work of L. E. Dickson.[2] Quite general results are obtainable by studying the form ϕ, without attempting to replace it by a form without cross-product terms. We shall consider mainly the case of positive c_i, though some of our results do not involve this restriction.

3. **Cases in which** (4) **implies** (1). If $ta - b^2$ is represented in ϕ for integers y_j, and x_i are obtained from (1_2) and (3), then $tx_1 = b + \sum c_jy_j$, and all the x_i are integers along with x_1. This proves

Theorem 1. *Let $tu \neq 0$. The number of solutions of* (1) *in integers x_i is equal to the number of solutions of* (4) *in integers y_j satisfying*

$$(6) \qquad c_2y_2 + \cdots + c_sy_s \equiv -b \pmod{t}.$$

Received October 7, 1940.

[1] G. Pall, Quarterly Journal of Mathematics, (Oxford), vol. 2(1931), pp. 136–143; to be referred to as QJ.

[2] L. E. Dickson, American Journal of Mathematics, vol. 56(1934), pp. 513–528. See also Dickson's *Modern Elementary Theory of Numbers*, Chicago, 1939, Chapter 10.

Now $\sum c_j y_j^2 \equiv \sum c_j y_j \equiv (\sum c_j y_j)^2 \pmod 2$. Hence

$$\phi = t \sum c_j y_j^2 - (\sum c_j y_j)^2 \equiv (t-1)(\sum c_j y_j)^2 \pmod{2t}. \tag{7}$$

Hence (4) and (6) imply the otherwise clearly necessary condition

$$a \equiv b \pmod 2. \tag{8}$$

Conversely, if (8) holds, (7) shows that all solutions of (4) satisfy

$$(\sum c_j y_j)^2 \equiv b^2 \pmod{\theta t}, \qquad \theta = 1 \ (t \text{ odd}),\ \theta = 2 \ (t \text{ even}). \tag{9}$$

For certain values of t, $K^2 \equiv b^2 \pmod{\theta t}$ implies

$$K \equiv b \equiv -b \pmod t, \quad \text{or} \quad \pm K \equiv b \not\equiv -b \pmod t. \tag{10}$$

Thus if $\pm t$ is 1 or 2, (10_1) holds for arbitrary b; if $\pm t$ is 4, (10_1) holds for b even, and (10_2) for b odd; if $\pm t$ is p or $2p$ (p an odd prime), (10_1) holds if $p \mid b$, (10_2) if $p \nmid b$. In certain cases $K^2 \equiv b^2 \pmod{\theta t}$ implies (10) for special forms of b. Call a prime-power p^n of type A or B in accordance with the following table:

$p > 2$: A if $n = 1$ and $p \mid b$; B if $n \geqq 1$ and $p \nmid b$;

$p = 2$: A if $n = 1$, b arbitrary, or if $n = 2$, b even; B if $n \geqq 2$ and $2 \mid b - 1$, or $n = 3$ and $4 \mid b - 2$, or $n = 4$ and $8 \mid b - 4$.

Set $\pm t = \prod p^n$ in powers of distinct primes. Then $K^2 \equiv b^2 \pmod{\theta t}$ implies (10_1) in cases (L), and (10_2) in cases (M), where

(L) every prime-power p^n in t is of type A;

(M) every p^n but one is of type A, and that one is of type B.

For other forms of t and b, the solvability of (4) and (8) may not imply that of (6). The reader may try some numerical examples in the table of §7.

THEOREM 2. *When ϕ belongs to a genus of one class, the numbers represented by ϕ are easily determined, and the cases* (L) *and* (M) *can be regarded as solved. If* (8) *holds, all solutions of* (4) *satisfy* (6) *in cases* (L), *and half of them satisfy* (6) *in cases* (M). *Case* (L) *or* (M) *holds for b arbitrary if $\pm t$ is* 1, 2, 4, p, *or* $2p$ *(p an odd prime).*

4. If $ta - b^2 = ta' - b'^2$ and $b = b' + ht$, then $a = a' + 2hb' + h^2 t$; and if (1) holds, then $a' = \sum c_i(x-h)^2$, $b' = \sum c_i(x-h)$. This proves again the theorem, also obvious from the preceding section:

THEOREM 3. *Equations* (1) *are solvable for all or none of the pairs* (a, b) *in a set* $(r, \rho)_t$; *where* $(r, \rho)_t$ *denotes the set of all pairs* (a, b) *such that*

$$ta - b^2 = r, \qquad b \equiv \pm\rho \pmod t. \tag{11}$$

For example, we cannot solve $19 = x_1^2 + \cdots + x_9^2$, $13 = x_1 + \cdots + x_9$, since $19t - 169 = 2$, and the set $(2, 4)_9$ contains $a = 2$, $b = 4$, with $a < b$.

The author proved in QJ for the case $c_1 = \cdots = c_s = 1$ (and this can be generalized) that if $s \geqq 5$ and (8) holds, the equations $a = \sum x_i^2$, $b = \sum x_i$ are solvable if and only if $sa - b^2 \geqq 0$, and there does not exist in the set $(r, \rho)_s$ of (a, b) a pair (a_0, b_0) with $a_0 < b_0$; and a simple method of obtaining the few unsolvable sets for any given s was derived. As a corollary we can now state the numbers represented by the forms $\phi_s = (s - 1) \sum y_j^2 - 2 \sum y_j y_k$:

Let Λ_s denote any non-negative integer of the form $sa - b^2$, $a \equiv b \pmod 2$. All numbers represented by $\phi_s = s(\sum y_j^2) - (\sum y_j)^2$ are of the form Λ_s. If $s \geqq 5$, ϕ_s represents every Λ_s, except that ϕ_9 does not represent 2; $\phi_{10} \neq 1, 4, 5$; $\phi_{11} \neq 2, 6, 8$; $\phi_{12} \neq 3, 8, 12$; $\phi_{16} \neq 7, 16, 23, 31$; and in general the largest exception does not exceed $[\frac{1}{4}(s^2 - 8s)]$.

5. **Properties of ϕ.** Since the "form" $b^2 + \phi$ is derived from $t \sum c_i x_i^2$ by the inverse of the transformation of determinant $(-)^{s-1}t$ defined by (1_2) and (3), $\det \phi = ut^{s-2}$, and the index of ϕ equals the number of negative terms tc_i. Hence ϕ is positive definite only if all the c_i have the same sign; and negative definite if $s - 1$ of the c_i are positive but $t = \sum c_i$ is negative, or vice-versa. These will be the difficult cases, since indefinite forms are much more frequently in genera of one class.

The effect of interchanging c_j and c_k, where $j, k \geqq 2$, is the same as that of interchanging y_j and y_k. The form obtained on interchanging c_1 and c_2 is the same as that obtained by the transformation

$$y_2 = Y_2, \; y_3 = Y_2 - Y_3, \; \cdots, \; y_s = Y_2 - Y_s. \tag{12}$$

LEMMA 1. *Permuting the c_i replaces ϕ by a properly or improperly equivalent form.*

It may be noted that $\pm \sum c_j y_j \pmod t$ remains unchanged.

Let $1 \leqq n \leqq s - 2$, $2 \leqq i_1 < i_2 < \cdots < i_n \leqq s$, $2 \leqq j_1 < j_2 < \cdots < j_n \leqq s$. The minor determinant Δ of the matrix of ϕ, in which the row positions are $i_1 - 1, \cdots, i_n - 1$ and column positions $j_1 - 1, \cdots, j_n - 1$, is zero (since two columns are proportional) unless either: (α) the i_ν and j_ν coincide, or (β) $n - 1$ of the i_ν are the same as $n - 1$ of the j_ν, but one index j_μ is distinct from all the i_ν. In the respective cases (α) and (β) we find that $\pm\Delta$ is

$$c_{i_1} \cdots c_{i_n} t^{n-1}(t - c_{i_1} - \cdots - c_{i_n}) \qquad \text{and} \qquad c_{i_1} \cdots c_{i_n} t^{n-1} c_{j_\mu}. \tag{13}$$

In particular, the form adjoint to ϕ is $t^{s-3}\psi$, where

$$\psi = \sum_{j}^{2,\cdots,s} (u/c_1 c_j)(c_1 + c_j) z_j^2 + 2(u/c_1) \sum_{j<k}^{2,\cdots,s} z_j z_k. \tag{14}$$

The expressions (13) show that the g.c.d. of the minor determinants of order n in the matrix of ϕ is $|\, t \,|^{n-1} \delta_n$, where

(15) δ_n is the g.c.d. of the products $n + 1$ at a time of $c_1, \cdots, c_s$.

Hence if $\phi^{(n)}$ is the n-th concomitant form of ϕ, the form F_n given by $\phi^{(n)} = |t|^{n-1}\delta_n F_n$ is properly or improperly primitive ($n = 1, \cdots, s-2$). To see which it is we can by Lemma 1 suppose the c_i arranged according to ascending powers of 2, say $2^{\theta_i} \| c_i$ ($i = 1, \cdots, s$), where $0 = \theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_s$. The least power of 2 dividing the cross-product coefficients 2Δ, with Δ as in (13_2), is that dividing $2c_2 \cdots c_{n+1}c_{n+2}t^{n-1}$. This power of 2 will be a factor of all the numbers (13_1) if and only if $t - c_2 - \cdots - c_{n+1}$ is even; that is, F_n is improperly primitive if and only if

(16) $$n = r-3, r-5, r-7, \cdots, \text{ but } n \geqq 1,$$

where r is defined by $c_1, \cdots, c_{r-1}$ odd, $c_r, \cdots, c_s$ even; $r = s+1$ if u is odd.

The preceding remarks determine the order invariants of ϕ in terms of the c_i. The Smith-Minkowski invariants I_n are given by

$$I_1 = |t|\,\delta_2/\delta_1^2,\ I_2 = \delta_1\delta_3/\delta_2^2,\ \cdots,\ I_{s-3} = \delta_{s-4}\delta_{s-2}/\delta_{s-3}^2,\ I_{s-2} = \delta_{s-3}\,|u|\,/\delta_{s-2}^2.$$

In particular, if $s = 4$,

(17) $$I_1 = \pm\Omega = |t|\,\delta_2/\delta_1^2, \qquad I_2 = \pm\Delta = |u|\,\delta_1/\delta_2^2,$$

and by (16), F_2 is i.p. if all four c_i are odd, F_1 is i.p. if three c_i are odd, both F_1 and F_2 are p.p. if only one or two c_i are odd.

The generic characters can also be specified in terms of the c_i, but it seems to be simpler to determine them directly in any given case.

The equality of cross-product coefficients in (14) is noteworthy.

LEMMA 2. *If r is distinct from $0, a_2, \cdots, a_s$, and the form*

(18) $$F = \sum_{j}^{2,\cdots,s} a_j z_j^2 + 2r \sum_{j<k}^{2,\cdots,s} z_j z_k$$

is of non-zero determinant, then the equations

(19) $$rc_1 = (a_j - r)c_j \qquad (j = 2, \cdots, s)$$

determine non-zero integers $c_1, \cdots, c_s$ such that $t = \sum c_i \neq 0$, and such that the adjoint of F is proportional to ϕ in (5).

For by (19), ψ in (14) is proportional to F, and the adjoint of ψ, which has non-zero determinant with F, is proportional to (5).

6. The property associated with (16) shows that it is not true if $s \geqq 5$ that *every* class of forms contains a form of type (18), whence its adjoint would be connected with a pair of equations (1). However, if $s = 4$ at least, it is possible that every class contains an infinitude of forms of type (18), and is therefore associated with an infinitude of sets $[c_1, \cdots, c_s]$. For example, $x^2 + y^2 + z^2$ has the equivalent forms and sets of c_i:

(1, 5, 9, 2, 2, 2), (1, 5, 41, −2, −2, −2), (1, 21, 45, 4, 4, 4),

(1, 17, 161, 4, 4, 4), ⋯ , [6, 14, 21, −42], [42, 258, 602, −903],

[204, 492, 2091, −2788], [156, 1884, 6123, −8164],

etc., all with $t = -1$; the divisor δ_1 of ϕ in these cases is respectively 42, 1806, 8364, 24492, and the condition of solvability is that $(a + b^2)/\delta_1$ shall be a non-negative integer not of the form $4^h(8k + 7)$. As another example, the equivalent forms $(2, 5, 9, -2, -2, -2)$ and $(3, 3, 10, -2, -2, -2)$ of determinant 10 lead to $[28, 44, 77, -154]$ with $t = -5$, and $[5, 12, 12, -30]$ with $t = -1$. Many sets of c_i of different signs are thus associated with negative definite forms in genera of one class; we note a few of the smaller ones: $[1, 1, 1, -h]$ ($h = 4, 5, 6, 7, 8$), $[1, 1, 2, -k]$ ($k = 5, 6, 7, 8$), $[1, 3, 3, -12]$, etc.

However, such sets of positive c_i are much rarer. We prove

THEOREM 4. *Let $s = 4$. With every class of ternary quadratic forms in which the Eisenstein-reduced form happens to be of the type*

$$(a_2, a_3, a_4, r, r, r), \qquad 0 < 2r \leqq a_2 \leqq a_3 \leqq a_4, \tag{20}$$

is associated as in Lemma 2 a unique set of positive relative-prime integers c_i. Conversely, if the c_i are positive the adjoint of ϕ has its reduced form of type (20).

For (19) yields positive c_i if and only if F or $-F$ has

$$0 < r < a_2, a_3, a_4. \tag{21}$$

Permuting the z_j we can take $a_2 \leqq a_3 \leqq a_4$. If $a_2 \geqq 2r$ we have (20). If $a_2 < 2r$ the transformation $z_2 = Z_2 + Z_3 + Z_4$, $z_3 = -Z_3$, $z_4 = -Z_4$ (i.e., permuting c_1 and c_2) replaces F by

$$F_2 = (a_2, a_2 + a_3 - 2r, a_2 + a_4 - 2r, a_2 - r, a_2 - r, a_2 - r), \tag{22}$$

where $2(a_2 - r) \leqq a_2$, $a_2 + a_3 - 2r$, and $a_2 + a_4 - 2r$, so that F_2 is of type (20). The uniqueness of the reduced form (20) proves that at most one set of positive c_i can arise from a given class, though possibly an infinitude with c_i of different signs.

7. Positive ternary forms in genera of one class. A search was made through B. W. Jones' table[3] of Eisenstein-reduced forms to find forms of type (20) in genera of one class. Besides these the forms arising from sets of positive c_i with $\sum c_i \leqq 17$ were tested, when their determinants exceeded 200, by means of Smith's formula[4] for the mass of a positive ternary genus, the number of automorphs being found from a list in Dickson's Studies.[5] The result was the table on page 178.

8. An example in a genus of two classes. Let $c_1 = 1$, $c_2 = 2 = c_3$, $c_4 = 4$; then $\phi = 2F$, $F = (4, 7, 7, -2, -1, -1)$; and $(1, 9, 18)$ is in the genus of F. We prove that if $3 \nmid b$ (cf. §3), (1) is solvable if and only if $9a - b^2$ is a non-negative even integer not of the form $4^{q+1}(8n + 7)$ nor equal to 2.

[3] B. W. Jones, Bulletin No. 97, National Research Council, 1935.

[4] H. J. S. Smith, *Collected Mathematical Papers*, vol. 1, p. 499.

[5] L. E. Dickson, *Studies in the Theory of Numbers*, Chicago, 1930, pp. 179–180.

CASES WITH $s = 4$, $c_i > 0$, ϕ IN A GENUS OF ONE CLASS

c_1	c_2	c_3	c_4	t	Necessary conditions for (1) to be solvable are (8), $ta - b^2 \geqq 0$, $\delta_1 \mid ta - b^2$, where δ_1 is defined in (15), and that $ta - b^2$ is not of the following forms:	Additional conditions to assure solvability
1	1	1	1	4	$4^{q+1}(8n + 7)$, $q \geqq 0$	none
1	1	1	2	5	$5^{2q+2}(5n + 2, 3)$	none
1	1	1	3	6	$3^{2q+1}(3n + 2)$	none
1	1	2	2	6	$4^{q+1}(8n + 7)$	none
1	1	1	4	7	$4^{q+1}(8n + 7)$	none
1	1	2	3	7	$7^{2q+2}(7n + 1, 2, 4)$	none
1	2	2	2	7	$4^{q}(16n + 14)$	none
1	1	1	5	8	$5^{2q+1}(5n + 1, 4)$	$4 \nmid b$
1	1	2	4	8	$4^{q+2}(16n + 14)$	$4 \nmid b$
1	1	3	3	8	$3^{2q}(3n + 2)$	$4 \nmid b$
1	2	2	3	8	$4^{q+2}(8n + 5)$	$4 \nmid b$
1	3	3	3	10	$5^{2q+2}(5n + 2, 3)$	none
1	1	4	6	12	$3^{2q+1}(3n + 1)$, $8(8n + 1)$	$2 \mid b$ or $3 \mid b$
1	1	5	5	12	$4^{q+1}(8n + 7)$, $5n + 2, 3$	$2 \mid b$ or $3 \mid b$
1	3	4	4	12	$4^{q+2}(8n + 5)$, $4(4n + 1, 2)$, $9n + 3$	$2 \mid b$ or $3 \mid b$
1	3	3	6	13	$13^{2q+2}(13n + 2, 5, 6, 7, 8, 11)$	none
1	1	3	9	14	$7^{2q+2}(7n + 1, 2, 4)$, $3n + 2$, $9n + 3$	none
1	1	1	12	15	$5^{2q+2}(5n + 2, 3)$, $3^{2q+1}(3n + 2)$, $4^{q+1}(8n + 5)$	$3 \mid b$ or $5 \mid b$
1	2	2	10	15	$4^{q}(16n + 6)$, $5(5n + 1, 4)$	$3 \mid b$ or $5 \mid b$
3	4	4	4	15	$3^{2q+1}(3n + 2)$, $5^{2q+2}(5n + 2, 3)$, $4^{q+1}(8n + 5)$	$3 \mid b$ or $5 \mid b$
1	3	3	9	16	$3^{2q+2}(3n + 2)$	b or $\frac{1}{4}b$ odd
1	3	6	6	16	$4^{q+2}(8n + 5)$	b or $\frac{1}{4}b$ odd
1	5	5	5	16	$4^{q+2}(8n + 3)$	b or $\frac{1}{4}b$ odd
1	5	5	10	21	$7^{2q+2}(7n + 1, 2, 4)$	$3 \mid b$ or $7 \mid b$
1	3	9	9	22	$3^{2q+1}(3n + 2)$	none
1	9	9	9	28	$4^{q+1}(8n + 7)$	$2 \mid b$ or $7 \mid b$
1	13	13	13	40	$5^{2q+2}(5n + 2, 3)$	$5 \mid b$, $4 \nmid b$
5	6	15	30	56	$4^{q+2}(8n + 1)$	$7 \mid b$, $4 \nmid b$
5	9	45	45	104	$13^{2q+2}(13n + 2, 5, 6, 7, 8, 11)$	$13 \mid b$, $4 \nmid b$

For, $F = X^2 + Y^2 + 2Z^2$, where

$$X = -x + 2y + z, \qquad Y = -x + y - 2z, \qquad Z = x + y - z,$$

whence

$$9x = -X - 3Y + 5Z, \qquad 3y = X + Z, \qquad 9z = 2X - 3Y - Z.$$

Now $N = \frac{1}{2}(9a - b^2)$ is represented in $X^2 + Y^2 + 2Z^2$ if $N \neq 4^q(16n + 14)$. We need only secure Y prime to 3, $X \equiv -Z \pmod 3$, and $X - 5Z \equiv \pm 3 \pmod 9$; then the sign of Y can be chosen to make x, y, z integers.

Permuting X, Y we can assume Y prime to 3. Then unless $X = Z = 0$ to begin with, $X^2 + 2Z^2 = 3^\alpha m$, $\alpha \geqq 1$, m prime to 3. By elementary properties of the form $X^2 + 2Z^2$, $m = r^2 + 2s^2$ with only one of r and s prime to 3. If $\alpha = 1$, define $X = r + 2s$, $Z = \pm(r - s)$, which are prime to 3, and choose the sign to make $X \equiv -Z \pmod 3$; then $X - 5Z \equiv \pm 3 \pmod 9$, since $X - 5Z \equiv 0 \pmod 9$ implies $X^2 + 2Z^2 \equiv 0 \pmod 9$. If $\alpha = 2$, define $X = 3r$, $Z = 3s$; then $X - 5Z \equiv (3s \text{ or } 3r) \equiv \pm 3 \pmod 9$. If $\alpha > 2$, $3^{\alpha-2}m$ is of the form $t^2 + 2u^2$ with tu prime to 3; set $X = 3t$, $Z = \pm 3u$, and choose the sign to make $t - 5u \equiv \pm 1 \pmod 3$.

There remains the case of a square: $N = Y^2$. If $Y = 1$ there is in fact no solution satisfying the desired conditions. If Y is even, we have $4w^2 = w^2 + w^2 + 2w^2$; if Y contains a prime factor $p = 8n + 1$ or $8n + 3$, set $Y = wp$, $p = r^2 + 2s^2$, r odd, $Y^2 = (r^2 - 2s^2)^2w^2 + 2(2rsw)^2$. Finally, if Y contains a prime factor $p = 8n + 5$ or $8n + 7$, we use the case $\lambda = 2$ of Theorem 4 in a recent article,[6] whence $p^2 = 2y_1^2 + x_2^2 + x_3^2$, $x_3 \equiv 4 \pmod 8$, x_2 odd.

9. **Solvability in non-negative integers.** Let k be an integer, $c_i > 0$ ($i = 1, \cdots, s$), $b \geqq t(1 - k)$. Then sufficient conditions for (1) to be solvable in integers $\geqq 1 - k$ are obtained by adjoining to conditions for solvability in integers the single inequality

$$b^2 + 2c_ikb + c_itk^2 > a(t - c_i) \quad \text{for the least of the } c_i. \tag{23_k}$$

In fact *all integral solutions of* (1) *must then satisfy* $x_i \geqq 1 - k$.

This result is a special case of (27) in QJ, but since it seems to have been overlooked, we give a simple proof. First let $k = 1$. Evidently if (23_1) holds with $c_i > 0$, it holds for larger c_i. By (2) with one c_i omitted,

$$(t - c_i)(a - c_ix_i^2) \geqq (b - c_ix_i)^2 \qquad (i = 1, \cdots, s). \tag{24}$$

Hence if any $x_i \leqq -1$, $(t - c_i)(a - c_i) \geqq (b + c_i)^2$, and (23_1) is contradicted. For any k, set $a' = a + 2(k - 1)b + (k - 1)^2t = \sum c_i(x_i + k - 1)^2$, $b' = b + t(k - 1) = \sum c_i(x_i + k - 1)$, where now $x_i + k - 1 \geqq 0$. Replacing a and b in (23_1) by a' and b' gives (23_k).

Let $c_1 = \cdots = c_s = 1$, $k = 1$. If $s = 4$, (23) gives the familiar $b^2 + 2b + 4$

[6] B. W. Jones and G. Pall, Acta Mathematica, vol. 70(1939), pp. 165–191.

$> 3a$, and if $s = 5$, the worse condition $b^2 + 2b + 5 > 4a$; while one may expect a better result for mere existence of non-negative solutions if $s = 5$. Taking $x_5 = 0$ or 1, and applying the result for $s = 4$ to $a - x_5^2$, $b - x_5$, the author obtained $b^2 \geqq 3a - 5$, for $s = 5$, 6, and 7 (QJ, p. 140). However, if a and b are even, we have the better result for $s = 4$, $3b^2 + 8b + 16 > 8a$ (QJ, p. 140); but now the condition $4a - b^2 \neq 4^q(8n + 7)$ is no longer vacuous, as in the case a and b odd. None the less we can apply this for $s = 5$ by proving that $4(a - x_5^2) - (b - x_5)^2 \neq 4^q(8n + 7)$ for one of $x_5 = 0, \cdots, 7$, and $b - x_5$ even. We can thus, if $s \geqq 5$, replace $b^2 \geqq 3a - 5$, if $b > 14$, by $3b^2 - 34b + 499 > 8a$; and we can no doubt improve this for $s \geqq 6$, by considering $a - x_5^2 - x_6^2$, etc.

The method of QJ (p. 141) can be used to prove Dickson's result[7] that if $s \geqq 8$ and $I(s) = [\frac{1}{4}(s^2 - 8s)]$, then $a = \sum x_i^2$, $b = \sum x_i$ are solvable in integers $\geqq 0$ if $b \geqq 0$, $a \equiv b \pmod 2$, and $3a - 5 \leqq b^2 < sa - I(s)$. For, assuming this result for $s - 1$, consider $\phi(x) = (s - 1)(a - x^2) - I(s - 1) - (b - x)^2$, whence $s\phi(x) = (s - 1)(sa - b^2 - I(s)) + (s - 1)I(s) - sI(s - 1) - (sx - b)^2$. The integer x nearest b/s makes $\phi(x) \geqq 0$ (this completing the induction) since, as Dickson observes (cf. p. 289 of the reference in footnote 7), $(s - 1)I(s) - sI(s - 1) = \sigma^2$, where $s = 2\sigma$ or $2\sigma + 1$.

As for the case $s = 4$, the constant $\lambda = 3$ in the inequality $(\lambda + 1)b^2 + 8kb + 16k^2 > 4\lambda a$ cannot be lowered much, in general; for example, no solution in non-negative x_i exists if $a = 347$, $b = 31$. By using transformations of $y_2^2 + y_3^2 + y_4^2$ with rational coefficients, the writer can reduce λ to 2 if $2 \mid n$ or $9 \mid n$, to 218/81 if $n \equiv 11 \pmod{24}$, to 32/27 if $162 \mid n$, to 11/9 if $6561 \mid n$, where $n = 4a - b^2$.

McGill University and The Institute for Advanced Study.

[7] L. E. Dickson, Quarterly Journal of Mathematics, (Oxford), vol. 5(1934), pp. 283–290, Theorem 11.

A THEOREM OF BOAS

By Norman Levinson

Boas has proved the following theorem.[1]

Theorem. *If the entire function $f(z)$ satisfies*

$$\limsup_{|z| \to \infty} \frac{1}{|z|} \log |f(z)| < \log 2 \tag{1}$$

and $f(z)$ is not a polynomial, an infinite number of derivatives of $f(z)$ are univalent in the unit circle $|z| \leqq 1$.

Here we shall give a direct and quite simple proof of this theorem. Incidentally, as has been pointed out to me by Boas, this also furnishes a simple proof for a theorem of Takenaka. Takenaka's theorem[2] states that *if every derivative of an entire function $f(z)$ has a zero inside or on the unit circle and if* (1) *holds, then $f(z)$ is a constant.*

Obviously, Takenaka's theorem is an immediate consequence of the above stated theorem of Boas.

We now turn to the proof of the theorem of Boas.[3] By a trivial change of variable it will suffice to show that

$$\limsup_{|z| \to \infty} \frac{1}{|z|} \log |f(z)| < 1 \tag{2}$$

implies that an infinite number of derivatives of $f(z)$ are univalent in $|z| < \log 2$.

Let the power series for $f(z)$ be

$$\sum_0^\infty a_n z^n.$$

From the Cauchy integral formula for a_n and from (2) it follows that

$$|a_n| \leqq \frac{e^{(1-\epsilon)R}}{R^n}$$

for large R. In particular, if $R = n$,

$$a_n = O\left(\frac{e^n e^{-\epsilon n}}{n^n}\right) = o\left(\frac{1}{n!}\right).$$

Received November 18, 1940.

[1] R. P. Boas, *Univalent derivatives of entire functions*, this Journal, vol. 6(1940), p. 719.

[2] S. Takenaka, *On the expansion of integral transcendental functions in generalized Taylor series*, Proc. Physico-Math. Soc. Japan, vol. 14(1932), pp. 529–542.

[3] **Added February 10, 1941.** In a letter to Boas dated January 1, 1941 Pólya communicated the same proof as is given here.

Or

$$n!a_n = o(1), \qquad n \to \infty.$$

Thus from the sequence $\{a_n n!\}$ there can be chosen an infinite subsequence $\{a_{n_k} n_k!\}$ such that

$$(3) \qquad | a_{n_k} n_k! | \geqq | a_n n! |, \qquad n \geqq n_k.$$

Consider

$$f^{(n_k-1)}(z) = \sum_{j=0}^{\infty} a_{n_k+j-1} \frac{(n_k + j - 1)!}{j!} z^j.$$

A function $\sum_0^\infty A_n z^n$ is univalent in $| z | < R$ if

$$| A_1 | \geqq \sum_{n=2}^{\infty} n | A_n | R^{n-1}.$$

Thus $f^{(n_k-1)}(z)$ is univalent for $| z | < R$, if

$$| a_{n_k} | n_k! \geqq \sum_{j=1}^{\infty} | a_{n_k+j} | \frac{(n_k + j)!}{(j + 1)!} (j + 1) R^j.$$

By (3) this certainly holds if

$$1 \geqq \sum_{j=1}^{\infty} \frac{R^j}{j!} = e^R - 1,$$

or if $R \leqq \log 2$. Thus $f^{(n_k-1)}(z)$ is univalent for $| z | < \log 2$.

Massachusetts Institute of Technology.

EXTENSION OF THE RANGE OF A DIFFERENTIABLE FUNCTION

By M. R. Hestenes

1. **Introduction.** In 1934 Whitney[1] showed that a function $f(x) = f(x_1, \cdots, x_n)$ of class C^m on a closed set A in Euclidean n-space E can be extended so as to be of class C^m on the whole space E. In fact he proved that $f(x)$ can be extended so as to be of class C^∞ on $E - A$. Using this result he showed further that the extension can be made so that $f(x)$ is analytic on $E - A$. In the present paper two methods of extending the range of differentiable functions are given. The first method (given in §3 below) is applicable only when m is finite and the boundary of A has suitable properties. It is, however, sufficiently general to be of interest, and the proof is relatively simple. The method used is a generalization of the reflection principle used by L. Lichtenstein[2] when $n = 3$ and $m = 1$. The second method (given in §§4 and 5 below) is essentially a modification of the one given by Whitney and is applicable to functions of class C^m (m finite or infinite) on an arbitrary closed set A. The details of the proof appear to be simpler than those of Whitney's. The extension is of class C^∞ on $E - A$ in this case.

2. **Notations and definitions.** In the following pages we shall use essentially the notations and terminology used by Whitney.[3] An n-tuple $k_1, \cdots, k_n$ of non-negative integers will be denoted by a single symbol k, and we write

$$k! = k_1!k_2! \cdots k_n!, \qquad \sigma_k = k_1 + \cdots + k_n, \qquad f_k(x) = f_{k_1 \cdots k_n}(x)$$

$$f_0(x) = f_{0 \cdots 0}(x), \qquad D_0 f(x) = f(x), \qquad D_k f(x) = \frac{\partial^{k_1 + \cdots + k_n}}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}} f(x).$$

By the symbol

$$P_m(x, x') = \sum_k \frac{f_k(x')}{k!}(x - x')^k \qquad (\sigma_k \leqq m) \tag{1}$$

will be meant the sum of all terms of the form

$$\frac{f_{k_1 \cdots k_n}(x')}{k_1! \cdots k_n!}(x_1 - x_1')^{k_1} \cdots (x_n - x_n')^{k_n}$$

for which $\sigma_k \leqq m$. We set

$$P_{m;k}(x, x') = D_k P_m(x, x'), \tag{2}$$

Received November 21, 1940.

[1] H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Transactions of the American Mathematical Society, vol. 36(1934), pp. 63–89.

[2] L. Lichtenstein, *Eine elementare Bemerkung zur reellen Analysis*, Mathematische Zeitschrift, vol. 30(1929), pp. 794–795.

[3] Loc. cit., p. 64.

where the differentiation is taken with respect to the variable x. We observe that

$$D_l P_{m;k}(x, x') = P_{m;k+l}(x, x'), \qquad P_{m;k}(x', x') = f_k(x'), \tag{3}$$

where $l = (l_1, \cdots, l_n)$ and $k + l = (k_1 + l_1, \cdots, k_n + l_n)$.

Let m be a non-negative integer. A function $f(x) = f_0(x)$ will be said to be *of class* C^m *on a set* A *in terms of functions* $f_k(x)$ $(\sigma_k \leqq m)$ if the functions $f_k(x)$ are defined on A and if for every point $x = a$ in A and every constant $\epsilon > 0$, there is a neighborhood N of $x = a$ such that for every pair of points x, x' in NA one has

$$| f_k(x) - P_{m;k}(x, x') | < \epsilon r^{m-\sigma_k}, \tag{4}$$

where r is the distance between the points x and x'. A function $f(x)$ will be said to be *of class* C^∞ *on* A *in terms of functions* $f_k(x)$ $(\sigma_k < \infty)$ if it is of class C^m on A in terms of the functions $f_k(x)$ $(\sigma_k \leqq m)$ for every integer m.

Let $f(x)$ be a function of class C^m on a set A in terms of functions $f_k(x)$ $(\sigma_k \leqq m)$. If $m > 0$, then for each integer $m' < m$ the function $f(x)$ is of class $C^{m'}$ on A in terms of the functions $f_k(x)$ $(\sigma_k \leqq m')$. Moreover, the functions $f_k(x)$ are continuous on A and at each interior point of A one has $D_k f(x) = f_k(x)$. It follows that if A is an open set or the closure of an open set, the functions $f_k(x)$ are uniquely determined by $f(x)$. A function $f(x)$ is of class C^m on an open set A if and only if it is continuous and has continuous derivatives of all orders $\leqq m$ on A, as can be seen with the help of the inequality (4) and Taylor's formula.

LEMMA 1. *Let* $f(x), f_k(x)$ $(\sigma_k \leqq m)$ *be continuous functions on* E *and let* A *be a closed subset of* E. *If in terms of the functions* $f_k(x)$ $(\sigma_k \leqq m)$ *the function* $f(x)$ *is of class* C^m *on each of the sets* A *and* $E - A$, *then* $f(x)$ *is of class* C^m *on* E *and* $D_k f(x) = f_k(x)$ *on* E.

It is sufficient to show that $f(x)$ is of class C' on E. In fact we can suppose that $f(x)$ is a function of a single variable x. Moreover, we can assume that $f_1(x) = 0$ since this can be brought about by replacing $f(x)$ by $f(x) - F(x)$, where $F'(x) = f_1(x)$. Let $x = a$ be a boundary point of A. Given a constant $\epsilon > 0$, let N be a δ-neighborhood of $x = a$ such that for each point x' in NA one has $| f(x') - f(a) | < \epsilon | x' - a |$. Consider a point x in $N - NA$. Let I be the largest open interval in $E - A$ containing x and let x' be the end point of I between x and $x = a$. We have $f(x') = f(x)$ and x' in NA. Hence

$$| f(x) - f(a) | = | f(x') - f(a) | < \epsilon | x' - a | < \epsilon | x - a |.$$

It follows that $f'(a)$ exists and is equal to zero. The function $f(x)$ is accordingly of class C' on E, as was to be proved.

Let $f(x)$ be a function of class C^m on a set A in terms of functions $f_k(x)$ $(\sigma_k \leqq m)$. A function $F(x)$ on an open set R will be called an *extension* of $f(x)$ on R

if it is of class C^m on R and $F(x) = f(x)$, $D_kF(x) = f_k(x)$ ($\sigma_k \leqq m$) on RA. We are interested, in particular, in extensions of $f(x)$ on the whole space E.

The following lemma will be useful in the proofs of Theorems 1 and 2 below. The proof of Theorem 3 below is, however, independent of this lemma. The lemma is also of interest in itself. The numbers m, q appearing in this result may be infinite.

LEMMA 2.[4] *Let $f(x)$ be a function of class C^m on a set A in terms of functions $f_k(x)$ ($\sigma_k \leqq m$) and let q be an integer not less than m. Suppose to each point $x = a$ in the closure of A there is a neighborhood N of $x = a$ and an extension $g(x)$ of $f(x)$ on N of class C^q on $N - NA$. Then there is an extension $F(x)$ of $f(x)$ on E of class C^q on $E - A$.*

Briefly this lemma states that a function is extensible over the whole space if and only if it is extensible locally. In order to prove this result let B be the closure of A. With each point of B there is associated a neighborhood N on which $f(x)$ has an extension of class C^q on $N - NA$. It follows that the set B can be covered by a finite or denumerable set of spheres S_i ($i = 1, 2, \cdots$) with centers x^i in B and radii r_i such that there is an extension $g_i(x)$ of $f(x)$ on the $(3r_i)$-neighborhood of x^i which is of class C^q at the points of this neighborhood not in A. Extend the definition of $g_i(x)$ so that $g_i(x) = f(x)$ on A and $g_i(x) = 0$ at the remaining points of E. Let S_i' be the sphere of radius $2r_i$ about x^i. The spheres S_i can be chosen so that each point of E is contained in at most a finite number of the spheres S_1', S_2', $\cdots$. Denote by $h_i(x)$ a function of class C^∞ on E such that $h_i(x) = 0$ on S_i and $h_i(x) = 1$ on $E - S_i'$. For example, we can take $h_i(x) = h(z/r_i)$, where z is the distance from x to x^i and $h(t)$ is a function of a single variable t having $h(t) = 0$ when $t \leqq 1$, $h(t) = 1$ when $t \geqq 2$ and $h(t) = \exp[\varphi(t)/(1 - t)]$, $\varphi(t) = \exp(t - 2)^{-1}$ on $1 < t < 2$. Set $H_i(x) = h_1h_2 \cdots h_{i-1}(1 - h_i)$. It is easily seen that $H_1 + \cdots + H_p = 1$ on S_p, $H_i = 0$ ($i > p$) on S_p and $H_p = 0$ on $E - S_p'$. From this last property it follows that at each point of E at most a finite number of the functions H_1, H_2, $\cdots$ is different from zero by virtue of the fact that this point is contained in at most a finite number of the spheres S_1', S_2', $\cdots$. Moreover, the product g_iH_i, being identically zero on $E - S_i'$, is of class C^m on E and of class C^q on $E - A$. The function $F(x)$ defined by the series

$$F(x) = g_1H_1 + g_2H_2 + \cdots + g_iH_i + \cdots$$

is therefore well defined on E, is of class C^m on E and of class C^q on $E - A$. At a point of A we have $g_i = f$ and $H_1 + H_2 + \cdots = 1$ and hence also $F(x) = f(x)$. In fact at a point x' of A in the sphere S_i' we have $D_kg_i = f_k$. From this fact and the relation $D_kH_1 + D_kH_2 + \cdots = 0$ on A it is readily seen that $D_kF(x) = f_k(x)$ on A. The function $F(x)$ is accordingly an extension of $f(x)$ and the lemma is established.

[4] Cf. E. J. McShane, *Necessary conditions in generalized-curve problems of the calculus of variations*, this Journal, vol. 7(1940), pp. 25–27.

3. Extension by a reflection principle. The results described in the present section are based on the following lemma which is established by the use of a generalized reflection.

LEMMA 3. *Let $f(y, z) = f(y_1, \cdots, y_{n-1}, z)$ be a function of class C^m (m finite) on a set of points (y, z) with $y = (y_1, \cdots, y_{n-1})$ on a set S and z on an interval of the form $0 \leqq z < e$. There exists a function $g(y, z)$ of class C^m at all points (y, z) with y on S and z on $-e < z < e$ which is identical with $f(y, z)$ when z is non-negative.*

Here the set S may be taken to be an open set plus some of its boundary points so that the concept of class is determined completely by the function itself. In order to prove the lemma let $a_1, \cdots, a_{m+1}$ be the solutions of the $m + 1$ linearly independent equations

$$(5) \quad (-1)^j a_1 + (-\tfrac{1}{2})^j a_2 + \cdots + (-1/m+1)^j a_{m+1} = 1 \quad (j = 0, 1, \cdots, m).$$

At each point (y, z) with y on S and z on $-e < z < e$ set $g(y, z) = f(y, z)$ when $z \geqq 0$ and set

$$(6) \quad g(y, z) = a_1 f(y, -z) + a_2 f(y, -z/2) + \cdots + a_{m+1} f(y, -z/m + 1)$$

when $z < 0$. It is easily seen with the help of equations (5) that the function so defined has the properties described in the lemma.

THEOREM 1. *Let A be a closed set in $(x_1, \cdots, x_n)$-space whose boundary B is a non-singular manifold of class C^q (q finite). Let $f(x)$ be a function of class C^m ($m \leqq q$) on A. There exists a function $F(x)$ of class C^m on E such that $F(x) = f(x)$ on A.*

For let x' be a point of B. Since B is a non-singular manifold of class C^q, it is representable near x' by functions

$$x_i = x_i(y_1, \cdots, y_{n-1}) \quad (-e \leqq y_j \leqq e; i = 1, \cdots, n; j = 1, \cdots, n-1)$$

of class C^q whose matrix $\| x_{iy_j} \|$ has rank $n - 1$. Let b_i be the direction cosines of the inner normal to B at x'. By the use of the equations $x_i = x_i(y) + b_i z$ one obtains a representation of a neighborhood of x' on which the hypotheses of Lemma 3 are satisfied. It follows from Lemma 3 that $f(x)$ has an extension $g(x)$ on a neighborhood of x'. The theorem now follows from Lemma 2.

The method used above can be modified so as to be applicable to the case when the boundary B of A is made up of non-singular manifolds with suitable edges. In particular it is applicable to the case in which to each point x' of B there is an integer h ($1 \leqq h \leqq n$) and a non-singular transformation

$$(7) \qquad x_i = x_i(t_1, \cdots, t_n) \qquad (i = 1, \cdots, n)$$

of class C^q of a neighborhood N of x' such that the points of B in N are the totality of points in N determined by equations (7) and the relations $t_j \geqq 0$ ($j \leqq h$), $t_1 t_2 \cdots t_h = 0$. We shall say that B is piecewise of class C^q in this case.

THEOREM 2. *The result described in Theorem 1 is true also when the boundary B of A is piecewise of class C^q in the sense just described.*

Again, by Lemma 2 we can restrict ourselves to a neighborhood N of a point x' in B. By virtue of the transformation (7) we can suppose that the points of B in N are given by relations of the form

$$0 \leqq x_i < e \ (i \leqq h), \qquad x_1 x_2 \cdots x_h = 0, \qquad -e < x_i < e \ (i > h).$$

Consider first the case in which the set NA consists of the points x for which

$$0 \leqq x_i < e \ (i \leqq h), \qquad -e < x_i < e \ (i > h). \tag{8}$$

By Lemma 3 with $z = x_h$ it is seen that $f(x)$ can be extended so as to be of class C^m on the set (8) with h replaced by $h - 1$ and hence over a neighborhood of x'. It remains to consider the case when the points of N determined by (8) belong to $N - NA$. In this case let A_i $(i \leqq h)$ be the set of points in N for which $x_i \leqq 0$. Then $NA = A_1 + \cdots + A_h$. We now define successively functions $g_1(x), \cdots, g_h(x)$ of class C^m on N such that $f = g_1 + \cdots + g_h$ on $A_1 + \cdots + A_h$. This is done as follows: Let $g_1(x) = f(x)$ on A_1 and extend $g_1(x)$ to be of class C^m on N as described in the proof of Lemma 3 with $z = -x_1$. Having defined the function g_{i-1} $(i > 1)$, set $g_i = f - g_1 - \cdots - g_{i-1}$ on A_i and extend g_i to be of class C^m on N by the formula (6) with $z = -x_i$ and $f = g_i$. Since $g_i = 0$ on $(A_1 + \cdots + A_{i-1})A_i$, its extension is identically zero on the set $A_1 + \cdots + A_{i-1}$ by virtue of formula (6). Hence $f = g_1 + \cdots + g_i$ on $A_1 + \cdots + A_i$. The function $g(x) = g_1(x) + \cdots + g_h(x)$ is therefore identical with $f(x)$ on the set $NA = A_1 + \cdots + A_h$ and is accordingly an extension of $f(x)$ on N. This completes the proof of Theorem 2.

4. **Four lemmas.** The proof of Theorem 3 below is based on Lemma 1 and four further lemmas, the first of which is the following:

LEMMA 4. *Let $f(x)$ be a function of class C^m (m finite) on a closed set A in terms of functions $f_k(x)$ and let $x = a$ be a point on the boundary of A. For every constant $\eta > 0$ there is a neighborhood N of $x = a$ such that if $x = b$ is a point of $N - NA$, one has at $x = b$ the inequalities*

$$| P_{m;k}(x, a') - P_{m;k}(x, a'') | < \eta e^{m-\sigma_k}, \qquad | P_{m;k}(x, a') - f_k(a) | < \eta \tag{9}$$

for every pair of points a', a'' of A in the $5e$-neighborhood of $x = b$, where $P_{m;k}(x, x')$ is defined by equation (2).

In order to prove this result we use the relations

$$P_{m;k}(x, a') - P_{m;k}(x, a'') = \sum_l [f_{k+l}(a') - P_{m;k+l}(a', a'')] \frac{(x - a')^l}{l!}, \tag{10}$$

where the sum is taken over all n-tuples l such that $0 \leqq \sigma_l \leqq m - \sigma_k$. The second member of this equation is the Taylor expansion of the first member about $x = a'$, as can be seen by the use of the relations (3). Let w be the

number of terms in the last sum. Given a constant $\eta > 0$, choose ϵ such that $10^m w\epsilon = \eta$ and let N' be a neighborhood of $x = a$ such that the inequality (4) holds for every pair of points x, x' in $N'A$. Choose a constant $\delta > 0$ such that the 6δ-neighborhood of $x = a$ is in N' and let $x = b$ be a point in the δ-neighborhood N of $x = a$ at a distance $e > 0$ from A. Then any two points a', a'' of A in the $5e$-neighborhood of $x = b$ are in N' and at a distance at most $10e$ apart. We have accordingly

$$| f_{k+l}(a') - P_{m;k+l}(a', a'') | < \epsilon(10e)^{m-\sigma_{k+l}}$$

by virtue of the inequality (4). By the use of this inequality together with equation (10) and the definitions of ϵ and w one obtains the first inequality (9). Since $P_{m;k}(a, a) = f_k(a)$, we can decrease N, if necessary, so that the second relation (9) holds.

LEMMA 5. *Let A be a closed subset of E. The set $E - A$ is the sum of a denumerable set of spheres $S_1, S_2, \cdots$ having the following properties: The center x^i of S_i is at a distance $e_i = 3r_i$ from A, where r_i is the radius of S_i. Let S_i' be the sphere of radius $2r_i$ about x^i. There is an integer q such that each point of E is interior to at most q of the spheres $S_1', S_2', \cdots$. If S_i' and S_p' have a point $x = b$ in common at a distance e from A, then*

$$(11) \qquad r_p/5 \leqq r_i \leqq 5r_p, \qquad e_p/5 \leqq e_i \leqq 5e_p, \qquad 3e/5 \leqq e_i \leqq 3e$$

and a point a^i of A at a distance e_i from x^i is in the $5e$-neighborhood of $x = b$.

In order to prove this result let C_e be the set of all points of E at a distance e from A. Denote by T_e a subset of C_e such that the distance between any two points of T_e is at least $e/9$ and such that C_e is in the $(e/9)$-neighborhood of T_e. The set T_e is at most denumerable. Moreover, if ϵ is on the range $e \leqq \epsilon \leqq 9e/8$, the set C_ϵ is in the $(e/3)$-neighborhood of T_e. It follows that the spheres of radius $e/3$ with centers in T_e, where e ranges over the positive, zero, and negative integral powers of 9/8, form a denumerable set of spheres $S_1, S_2, \cdots$ whose sum is $E - A$. Let S_i' be the sphere related to S_i as described in the lemma. In order to show that the number of spheres in the set $S_1', S_2', \cdots$ having a common point is bounded, we observe first that if a point $x = b$ in C_e is in the $(2\epsilon/3)$-neighborhood of T_ϵ, the inequality $3e/5 \leqq \epsilon \leqq 3e$ holds and the distances between any two points of T_ϵ is at least $e/15$. From these results we see that if the spheres S_i' and S_p' contain $x = b$, their centers lie in the $3e$-neighborhood of $x = b$ and are at distance of at least $e/15$ apart. The number of spheres in the sequence $S_1', S_2', \cdots$ containing $x = b$ therefore cannot exceed an integral upper bound q of the number of points in the $3e$-neighborhood of $x = b$, every pair of which are at a distance of at least $e/15$ apart. This upper bound q is clearly independent of our choice of e and $x = b$. The last statement in the lemma is easily verified by the use of the relation $e_i = 3r_i$.

LEMMA 6. *Let $f(x)$ be a function of class C^m (m finite or infinite) on a closed set A in terms of functions $f_k(x)$ ($\sigma_k \leqq m$) and let S_i, S_i' ($i = 1, 2, \cdots$) be spheres*

related to A as described in the last lemma. There exist polynomials $G_i(x)$ $(i = 1, 2, \cdots)$ having the following property: Given a point $x = a$ on the boundary of A, a constant $\eta > 0$ and an integer $m' \leqq m$, there exists a neighborhood N of A such that at each point $x = b$ of N at a distance $e > 0$ from A the inequalities

$$(12) \qquad | D_k G_i(x) - D_k G_p(x) | < \eta e^{m' - \sigma_k}, \qquad | D_k G_p(x) - f_k(a) | < \eta$$

hold whenever the spheres S_i' and S_p' contain $x = b$.

In order to prove this result let a^i be a nearest point to the center x^i of S_i. If m is finite, the functions $G_i(x) = P_m(x, a^i)$, where $P_m(x, x')$ is given by equation (1), have the property described in the lemma. For if S_i' and S_p' contain a common point $x = b$ at a distance e from A, the points a^i and a^p are in the $5e$-neighborhood of $x = b$, by Lemma 5. By the use of Lemma 4 it is seen that the inequalities (12) are a consequence of the inequalities (9), provided that $e < 1$ as we can suppose.

Consider next the case when m is infinite and let R_β $(\beta = 1, 2, \cdots)$ be the sphere of radius β about a fixed point of A. Choose constants N_β $(\beta = 1, 2, \cdots)$ such that at each point x of AR_β the inequalities

$$(13) \qquad \left| \sum_k \frac{f_k(x)}{k!} z^k \right| < N_\beta \quad (\sigma_k = \beta), \qquad N_\beta \leqq N_\gamma \quad (\beta < \gamma)$$

hold for every n-tuple $z = (z_1, \cdots, z_n)$ having $z_1^2 + \cdots + z_n^2 = 1$. Select positive constants $\delta_\beta < 1$ satisfying the relations $N_\beta \delta_\beta < 1$, $2\delta_{\beta+1} < \delta_\beta$. Set $m_i = \beta$ when the distance e_i of the center of S_i satisfies the inequality $\delta_{\beta+1} \leqq e_i < \delta_\beta$ and set $m_i = 0$ when $e_i \geqq \delta_1$. The functions $G_i(x) = P_{m_i}(x, a^i)$, where $P_m(x, x')$ is given by (1), have the property described in the lemma. To prove this statement let $x = a$ be a point on the boundary of A, let η be a positive constant and m' be an integer. Choose an integer $\alpha > m'$ so that $x = a$ is in $R_{\alpha-1}$. By Lemma 4 with $m = \alpha$ there exists a δ-neighborhood N of $x = a$ such that if $x = b$ is a point of N at a distance $e > 0$ from A

$$(14) \quad | P_{\alpha;k}(x, a') - P_{\alpha;k}(x, a'') | < \tfrac{1}{2}\eta e^{\alpha - \sigma_k}, \qquad | P_{\alpha;k}(x, a') - f_k(a) | < \tfrac{1}{2}\eta$$

hold at $x = b$ for every pair of points a', a'' of A in the $5e$-neighborhood of $x = b$, provided that $\sigma_k \leqq \alpha$. Consider spheres S_i', S_p' containing $x = b$. By Lemma 5 the points a^i, a^p are in the $5e$-neighborhood of $x = b$. The inequalities (14) therefore hold when $a' = a^p$ and $a'' = a^i$. Restrict the δ-neighborhood N of $x = a$ so that $5\delta < \delta_\alpha$, $5\delta < \frac{1}{2}$, $5^\alpha \cdot 72\delta < \eta$. For $j = i$ or p we then have, by (11), the inequality $e_j \leqq 3e < 3\delta < \delta_\alpha$. Hence $m_j \geqq \alpha$. Setting $t = 5e$ we have also, by (11), $t \leqq 9e_j$ and $t < 5\delta < \frac{1}{2}$. Consequently for $j = i$ or p we have

$$\frac{N_{m_j} t^{\alpha+1}}{1 - t} < 18 e_j N_{m_j} t^\alpha < 18 t^\alpha = 18(5e)^\alpha < \tfrac{1}{4}\eta e^{\alpha-1}$$

...

since $5^{\alpha}\cdot 72e < 5^{\alpha}\cdot 72\delta < \eta$ by virtue of our choice of δ. By the use of these inequalities, the relations (13) and the fact that a^j is in the t-neighborhood of $x = b$ we find that the inequality

$$(15) \qquad | P_{m_j;k}(x, a^j) - P_{\alpha;k}(x, a^j) | < N_{m_j}(t^{\alpha+1} + \cdots + t^{m_j}) < \frac{N_{m_j} t^{\alpha+1}}{1 - t} < \tfrac{1}{4}\eta e^{\alpha-1}$$

holds at $x = b$ when $j = i$ or p. The relations (12) are now easy consequences of the inequalities (14) with $a' = a^p$, and $a'' = a^i$, the inequality (15) with $j = i$ and $j = p$ and the relations $e < 1$, $m' < \alpha$ and $D_k G_j(x) = P_{m_j;k}(x, a^j)$.

LEMMA 7. *Let S_i, S_i' $(i = 1, 2, \cdots)$ be related to a closed set A as described in Lemma 5. There exist functions $H_i(x)$ $(i = 1, 2, \cdots)$ of class C^∞ on E such that for every integer p one has $H_1 + \cdots + H_p = 1$ on S_p, $H_i = 0$ $(i > p)$ on S_p, $H_p = 0$ on $E - S_p'$ and*

$$(16) \qquad | D_k H_p(x) | < \frac{M_k}{e^{\sigma_k}}$$

on E, where M_k is a positive constant independent of p and e is the distance from x to A. At most q of the functions $H_1, H_2, \cdots$ are different from zero at any point of E, where q is defined as in Lemma 5.

In order to define the functions $H_i(x)$ let $h(x)$ be a function of class C^∞ having $h = 0$ on the unit sphere about the origin and $h = 1$ exterior to the sphere of radius 2 about the origin. A function of this type was constructed in the proof of Lemma 2 above. Let U_k be an upper bound of the value $| D_k h(x) |$ on E. We can suppose that $U_k \leqq U_l$ when $\sigma_k \leqq \sigma_l$. Set

$$h_i(x) = h\left(\frac{x_1 - x_1^i}{r_i}, \cdots, \frac{x_n - x_n^i}{r_i}\right),$$

where x^i is the center of S_i and r_i is its radius. Then $h_i = 0$ on S_i, $h_i = 1$ on $E - S_i'$ and $U_k/r_i^{\sigma_k}$ is an upper bound for $| D_k h_i(x) |$ on E. Since $D_k h_i = 0$ on $E - S_i'$ and the distance e of x from A satisfies the relation $e < 5r_i$ when x is in S_i', we have

$$(17) \qquad | D_k h_i(x) | < \frac{5^{\sigma_k} U_k}{e^{\sigma_k}}$$

on E. Set $H_1 = 1 - h_1$ and $H_p = h_1 h_2 \cdots h_{p-1}(1 - h_p)$ $(p > 1)$. The function H_p is of class C^∞ on E and is identically zero on $E - S_p'$ and on each S_i when $i > p$. Moreover, $H_1 + \cdots + H_p = 1 - h_1 \cdots h_p = 1$ on S_p. Finally the relation (16) holds with $M_k = (5q)^{\sigma_k} U_k^q$, where q is defined as in Lemma 5. For let x' be a point in S_p' and let $S_{i_1}', \cdots, S_{i_s}'$ be the spheres S_i' $(i < p)$ containing x'. Then $s < q$ and $H_i = h_{i_1} \cdots h_{i_s}(1 - h_p)$. The derivative $D_k H_i$ is the sum of $(s + 1)^{\sigma_k} \leqq q^{\sigma_k}$ terms of the form

$$(D_{k_1} h_{i_1}) \cdots (D_{k_s} h_{i_s})(D_{k_{s+1}}(1 - h_p)) \qquad (k_1 + \cdots + k_{s+1} = k),$$

each term counted a number of times equal to its multiplicity. The absolute value of this term is less than the value $5^{\sigma_k} U_k^q / e^{\sigma_k}$, by (17). The sum of $(s + 1)^{\sigma_k}$ such terms is accordingly exceeded by M_k/e^{σ_k}, where M_k is defined as above. This proves the inequality (16). The last statement in the lemma follows from the fact that each point of E is interior to at most q of the spheres S_1', S_2', $\cdots$ and the relation $H_i = 0$ on $E - S_i'$.

5. Extension of functions defined on an arbitrary closed set. The result due to Whitney[5] which we shall prove is given in the following

THEOREM 3. *Let $f(x)$ be a function of class C^m (m finite or infinite) on a closed set A in terms of functions $f_k(x)$ ($\sigma_k \leqq m$). There exists a function $F(x)$ of class C^m on E and of class C^∞ on $E - A$ such that $F(x) = f(x)$, $D_k F(x) = f_k(x)$ ($\sigma_k \leqq m$) on A.*

In order to establish this result let S_i, S_i' ($i = 1, 2, \cdots$) be spheres and $G_i(x)$, $H_i(x)$ ($i = 1, 2, \cdots$) be functions related to A and to $f(x)$ as described in Lemmas 5, 6 and 7. The function $F(x)$ that is identical with $f(x)$ on A and is defined by the series

$$F(x) = G_1(x)H_1(x) + \cdots + G_i(x)H_i(x) + \cdots \tag{18}$$

on $E - A$ has the properties described in the theorem. It is well defined on $E - A$ since at most q of its terms are different from zero at a point of $E - A$, by virtue of the last statement in Lemma 7. Moreover, it is of class C^∞ on $E - A$ since each of its terms is of class C^∞ on $E - A$. It remains to show that $F(x)$ is of class C^m on E. To this end we observe first that, in view of the properties of the functions H_i described in Lemma 7, the equations

$$F(x) = \sum_{i=1}^{p} G_i H_i = G_p + \sum_{i=1}^{p} (G_i - G_p) H_i$$

and hence also the equation

$$D_k F(x) - D_k G_p(x) = \sum_{i=1}^{p} D_k[(G_i - G_p)H_i] \tag{19}$$

hold on the sphere S_p. Consider now a point $x = a$ on the boundary of A and choose an integer $m' \leqq m$. Let M be the maximum of the numbers M_k ($\sigma_k \leqq m'$) described in Lemma 7. Given a constant $\epsilon > 0$, choose η such that $2q2^{m'} M\eta = \epsilon$, where q is defined as in Lemma 5. Let N be a δ-neighborhood ($\delta < 1$) of $x = a$ that is related to η and m' as described in Lemma 6. Consider a point $x = b$ in N at a distance $e > 0$ from A and let S_p be one of the spheres S_1, S_2, $\cdots$ containing $x = b$. By Lemma 6 the inequality

$$| D_k G_p(x) - f_k(a) | < \eta < \tfrac{1}{2}\epsilon \qquad (\sigma_k \leqq m') \tag{20}$$

[5] Loc. cit., p. 69 (Lemma 2).

holds at $x = b$. Moreover, if the sphere S_i' also contains $x = b$, we have, by virtue of the inequalities (12) and (16), the further relation

$$| D_l(G_i - G_p) | \, | D_{k-l}H_i | < \eta M_{k-l} e^{m'-\sigma_k} \leqq \eta M \qquad (\sigma_l \leqq \sigma_k \leqq m')$$

at $x = b$ since $e < \delta < 1$. The derivative $D_k[(G_i - G_p)H_i]$ $(\sigma_k \leqq m')$ is the sum of $2^{\sigma_k} \leqq 2^{m'}$ terms of the form $(D_lG_i - D_lG_p)D_{k-l}H_i$. We have accordingly

$$(21) \qquad | D_k[(G_i - G_p)H_i] | \leqq 2^{m'} M\eta = \frac{\epsilon}{2q} \qquad (\sigma_k \leqq m')$$

at $x = b$. Now the i-th term in the second member of equation (19) is different from zero at $x = b$ only if S_i' contains $x = b$. Moreover, at most q of these spheres contain $x = b$, by Lemma 5. Using this fact we see from (19) and (21) that

$$| D_kF(x) - D_kG_p(x) | < \tfrac{1}{2}\epsilon \qquad (\sigma_k \leqq m').$$

Combining this result with (20) we find that the inequality

$$| D_kF(x) - f_k(a) | < \epsilon \qquad (\sigma_k \leqq m')$$

holds on $N - NA$. The function $F_k(x)$ $(\sigma_k \leqq m)$ that is identical with $f_k(x)$ on A and identical with $D_kF(x)$ on $E - A$ is therefore continuous at $x = a$ and hence on E. It follows from Lemma 1 that the function $F(x)$ is of class C^m on E and has $D_kF(x) = F_k(x) = f_k(x)$ on A. This completes the proof of Theorem 3.

THE UNIVERSITY OF CHICAGO.

FORMULATIONS OF THE HAUSDORFF INCLUSION PROBLEM

By H. L. Garabedian, Einar Hille and H. S. Wall

1. **Introduction.** A sequence $\{x_n\}$ of complex numbers is called a *regular sequence* if there exists a *regular mass function* $\phi_x(u)$ such that[1]

$$x_n = \int_0^1 u^n \, d\phi_x(u) \qquad (n = 0, 1, 2, \cdots).$$

The conditions on $\phi_x(u)$ are that it shall be of bounded variation on the interval $0 \leqq u \leqq 1$, continuous at $u = 0$, and that

$$\phi_x(u) = \begin{cases} 0 & \text{if } u \leqq 0, \\ 1 & \text{if } u \geqq 1, \\ \frac{1}{2}[\phi_x(u-0) + \phi_x(u+0)] & \text{if } 0 \leqq u < 1. \end{cases}$$

With each regular sequence $\{x_n\}$ there is associated a *regular moment function*

$$x(z) = \int_0^1 u^z \, d\phi_x(u),$$

and a *regular moment generating function*

$$f_x(t) = x_0 - x_1 t + x_2 t^2 - \cdots = \int_0^1 \frac{d\phi_x(u)}{1 + tu}.$$

If $\{x_n\}$ is a regular sequence, the corresponding *Hausdorff transform* of a sequence $\{s_n\}$ is given by

$$t_m = \sum_{n=0}^{n=m} C_{m,n} \Delta^{m-n} x_n \cdot s_n \qquad (m = 0, 1, 2, \cdots),$$

where $C_{m,n} = m!/n!(m-n)!$, and $\Delta^i x_j = x_j - C_{i,1} x_{j+1} + C_{i,2} x_{j+2} - \cdots$. This defines a regular *Hausdorff method of summation* which is denoted by the symbol $[H, \phi_x(u)]$. Let $\{a_n\}$ and $\{b_n\}$ be two regular sequences and $[H, \phi_a(u)]$, $[H, \phi_b(u)]$ the corresponding Hausdorff methods of summation. If $b_n \neq 0$ $(n = 0, 1, 2, \cdots)$, Hausdorff showed that $[H, \phi_a(u)] \supset [H, \phi_b(u)]$ if and only if the sequence $\{a_n\}$ is *divisible* by the sequence $\{b_n\}$; i.e., $a_n = b_n c_n$ $(n = 0, 1, 2, \cdots)$,

Received November 22, 1940; presented to the American Mathematical Society, September 10, 1940.

[1] The Stieltjes integrals discussed in this paper may always be taken in the Riemann-Stieltjes sense, but in §§3, 4 much is gained by using the Lebesgue definition instead. When integrals are to be taken in the Lebesgue-Stieltjes sense, we state so explicitly.

where $\{c_n\}$ is a regular sequence. By the *Hausdorff inclusion problem* we understand the problem of determining whether or not one regular sequence is divisible by a second regular sequence.

Hille and Tamarkin [4] stated (without proof) that either of the following conditions is necessary and sufficient for $\{a_n\}$ to be divisible by $\{b_n\}$.

(i) *There exists a regular moment function $c(z)$ such that*

$$a(z) \equiv b(z)c(z) \quad \text{if} \quad \Re(z) \geqq 0.$$

(ii) *There exists a regular mass function $\phi_c(u)$ such that*[2]

$$\phi_a(v) = \int_0^1 \phi_b(v/u)\, d\phi_c(u)$$

for all except at most a countable set of values of v in the interval $0 \leqq v \leqq 1$.

In this paper we obtain a third condition equivalent to these, namely:

(iii) *There exists a regular mass function $\phi_c(u)$ such that*

$$f_a(t) = \int_0^1 f_b(tu)\, d\phi_c(u) \quad \text{if} \quad |t| < 1.$$

We prove that $\{a_n\}$ is divisible by $\{b_n\}$ if and only if (iii) holds; and then prove that these formulations are equivalent.

In case $\phi_b(u)$ is the regular mass function for Cesàro summability (C, α), the integral equation in (ii) can be thrown into the form of an Abel integral equation. From this we obtain a set of necessary and sufficient conditions upon $\phi_a(u)$ in order that $[H, \phi_a(u)] \supset (C, \alpha)$, when α is an integer $\geqq 1$.

The latter part of the paper is concerned with applications of the preceding to the following problem. Let $\{a_n\}$ and $\{b_n\}$ be regular sequences defining equivalent Hausdorff methods of summation; to determine whether or not corresponding *row* sequences in the difference matrices $(\Delta^m a_n)$ and $(\Delta^m b_n)$ define equivalent Hausdorff methods. We have given a partial answer to this question in the case that the given sequences are the sequences defining the Cesàro and Hölder methods of summability of order α.

2. Formulation of the Hausdorff inclusion problem as a generalized moment problem. By the moment problem for the interval (0, 1) is ordinarily understood the problem of determining a bounded monotone function or a function of bounded variation $\phi_c(u)$ $(0 \leqq u \leqq 1)$ such that

$$a_n = \int_0^1 u^n\, d\phi_c(u) \qquad (n = 0, 1, 2, \cdots),$$

[2] This integral equation was used by R. Schmidt [12] in his doctoral dissertation.—Hille and Tamarkin state merely that the exceptional set is of measure zero. As will be brought out below, the structure of the exceptional set depends essentially upon the interpretation given to the integral.

the sequence $\{a_n\}$ having been given. In other terms $\phi_c(u)$ must satisfy the equation

$$f_a(t) = \int_0^1 f_b(ut)\, d\phi_c(u),$$

when $f_b(t) = 1/(1 + t)$, and $f_a(t) = a_0 - a_1 t + a_2 t^2 - \cdots$ is a given power series. We may generalize this at once and state the Hausdorff inclusion problem in terms of the moment generating functions $f_a(t), f_b(t)$ as follows:

THEOREM 2.1. *The regular sequence $\{a_n\}$ is divisible by the regular sequence $\{b_n\}$ if and only if there exists a regular mass function $\phi_c(u)$ such that*

$$f_a(t) = \int_0^1 f_b(ut)\, d\phi_c(u), \qquad |t| < 1, \tag{2.1}$$

where $f_a(t)$, $f_b(t)$ are the regular moment generating functions associated with $\{a_n\}$, $\{b_n\}$.

To prove this, suppose first that

$$a_n = b_n c_n \qquad (n = 0, 1, 2, \cdots), \tag{2.2}$$

where $\{c_n\}$ is a regular sequence, and consider the expression

$$\int_0^1 \int_0^1 \frac{d\phi_b(u)}{1 + uvt}\, d\phi_c(v), \qquad |t| < 1, \tag{2.3}$$

where $\phi_c(v)$ is the regular mass function associated with $\{c_n\}$. Inasmuch as $\int_0^1 d\phi_b(u)/(1 + uvt)$ is a continuous function of v $(0 \leqq v \leqq 1)$ for any t such that $|t| < 1$, it follows that the integral (2.3) exists. Moreover,

$$\int_0^1 \frac{d\phi_b(u)}{1 + uvt} = b_0 - b_1 tv + b_2 t^2 v^2 - \cdots,$$

where the series on the right converges uniformly for $0 \leqq v \leqq 1$ for any fixed value of t such that $|t| < 1$. We may therefore integrate this series term by term over the interval $0 \leqq v \leqq 1$, and get for (2.3) the series expansion $\sum_{i=0}^{i=\infty} b_i c_i (-t)^i$ which, by (2.2), is equal to $\sum_{i=0}^{i=\infty} a_i (-t)^i$, which is $f_a(t)$. Hence (2.1) holds.

Conversely, if (2.1) holds for some regular mass function $\phi_c(u)$, we see at once that (2.2) holds with $c_n = \int_0^1 u^n\, d\phi_c(u)$ $(n = 0, 1, 2, \cdots)$. It is clear that if the sequence $\{a_n\}$ is divisible by the sequence $\{b_n\}$, then $\{a_n\}$ is also divisible by $\{a_n/b_n\}$. This fact can be stated as follows.

THEOREM 2.2. *The regular sequence $\{a_n\}$ is divisible by the regular sequence $\{b_n\}$ if and only if there exists a regular moment generating function $f_c(t)$ such that*

$$f_a(t) = \int_0^1 f_c(tu)\, d\phi_b(u), \qquad |t| < 1, \tag{2.4}$$

where $f_a(t)$ is the regular moment generating function associated with $\{a_n\}$ and $\phi_b(u)$ the regular mass function associated with $\{b_n\}$.

The function $f_a(t)$ is an analytic function of t, holomorphic in the domain T obtained by deleting from the complex t-plane the line segment from $-\infty$ to -1 of the real axis. The same is true of $f_b(t)$ and $f_c(t)$. From this remark we see at once that formulas (2.1) and (2.4) are valid for t in T and not merely in $|t| < 1$.

Since $\phi_a(u) = 0$ for $u \leqq 0$ and equals 1 for $1 \leqq u$, we have for t in T and any ω $(1 \leqq \omega \leqq \infty)$

$$f_a(t) = \int_{-\infty}^{\omega} \frac{d\phi_a(u)}{1 + tu} = \frac{1}{1 + t\omega} + t\int_{-\infty}^{\omega} \frac{\phi_a(u)}{(1 + tu)^2}\, du.$$

Letting $\omega \to +\infty$, we get for $t \neq 0$

$$f_a(t) = t\int_{-\infty}^{\infty} \frac{\phi_a(u)}{(1 + tu)^2}\, du. \tag{2.5}$$

This alternative representation of $f_a(t)$ will be useful in a later discussion.

3. Formulation of the inclusion problem in terms of the integral equation of R. Schmidt. This formulation, save for the determination of the exceptional set, is due to Hille and Tamarkin [4], and may be stated as follows.

THEOREM 3.1. *The regular sequence $\{a_n\}$ with regular mass function $\phi_a(u)$ is divisible by the regular sequence $\{b_n\}$ with regular mass function $\phi_b(u)$ if and only if there exists a regular mass function $\phi_c(u)$ such that*

$$\phi_a(u) = \int_0^1 \phi_c(u/v)\, d\phi_b(v), \tag{3.1}$$

$$\phi_a(u) = \int_0^1 \phi_b(u/v)\, d\phi_c(v), \tag{3.2}$$

for all except at most a countable set of values of u in the interval $0 < u < 1$. Here the integrals are taken in the sense of Lebesgue-Stieltjes and exist for all values of u. If $\phi_c(1 - 0) = \phi_c(1)$, equation (3.1) holds for all values of u; otherwise the points of discontinuity of $\phi_b(u)$ in $0 < u < 1$ form the exceptional set. The exceptional set of the equation (3.2) is obtained by interchanging the subscripts b and c in the preceding sentence.

It should be observed that

$$\phi_b(1) - \phi_b(1 - 0) = \lim_{n\to\infty} b_n \qquad \text{and} \qquad \phi_c(1) - \phi_c(1 - 0) = \lim_{n\to\infty} a_n/b_n$$

if $\{a_n\}$ is divisible by $\{b_n\}$, so that the presence or absence of exceptional sets can be decided a priori. The exceptional set of (3.1) is of course given as soon as $\phi_b(u)$ is known, but that of (3.2) is determined only implicitly by the original sequences.

The integrals in (3.1) and (3.2) exist for all values of u and define functions of bounded variation, but these functions are not necessarily regular everywhere. The points in $0 < u < 1$ where the condition $\phi(u) = \frac{1}{2}[\phi(u+0) + \phi(u-0)]$ does not hold form the exceptional set for the relation in question. Outside of the exceptional sets the integrals agree with each other and with $\phi_a(u)$.

A different situation arises when the integrals are taken in the sense of Riemann-Stieltjes. Here the integrals do not exist for all values of u, but when they do exist, they are also regular. The exceptional sets are now the values of u for which the integrals do not exist. Thus the exceptional set associated with (3.1) consists of all points of the form $u_{b,m} \cdot u_{c,n}$, where $u_{b,m}$ is any point of discontinuity of $\phi_b(u)$ in $0 < u \leqq 1$ and $u_{c,n}$ any point of discontinuity of $\phi_c(u)$ in $0 < u < 1$. The exceptional set of (3.2) is obtained by interchanging the subscripts b and c in the preceding sentence.

We shall need two lemmas.

LEMMA 3.1. *Let $\phi_j(u)$ $(j = 1, 2)$ satisfy the following conditions:* (1) *$\phi_j(u)$ is of bounded variation in $-\infty < u < +\infty$,* (2) *$\phi_j(u) = 0$ for $u \leqq 0$ and continuous at $u = 0$,* (3) *$\phi_j(u) = \phi_j(1)$ for $1 \leqq u$, and* (4) *$\phi_j(u)$ is regular for $0 \leqq u < 1$, i.e., $\phi_j(u) = \frac{1}{2}[\phi_j(u+0) + \phi_j(u-0)]$. Then the Lebesgue-Stieltjes integral*

$$\phi_3(u) = \int_0^1 \phi_1(u/v)\, d\phi_2(v) \tag{3.3}$$

exists for all values of u and $\phi_3(u)$ satisfies conditions (1)–(3). *If $\phi_1(1-0) = \phi_1(1)$, condition* (4) *is also satisfied for all values of u $(0 \leqq u < 1)$, but if $\phi_1(1-0) \neq \phi_1(1)$, then* (4) *holds except at the points of discontinuity of $\phi_2(u)$ in $0 < u < 1$.*

In proving the lemma we may assume that $\phi_1(u)$ and $\phi_2(u)$ are real monotone non-decreasing functions since every function of bounded variation is a linear combination with coefficients ± 1, $\pm i$ of real monotone non-decreasing functions and $\phi_3(u)$ is a bilinear functional of $\phi_1(u)$ and $\phi_2(u)$.

A function of bounded variation being bounded and measurable with respect to any other function of bounded variation, condition (1) implies the existence of $\phi_3(u)$ for all values of u. Assuming $\phi_1(u)$ and $\phi_2(u)$ to be real monotone non-decreasing functions, we see that $\phi_3(u)$ has the same property. Further $\phi_3(u) = 0$ for $u \leqq 0$ and $\phi_3(u) = \phi_3(1) = \phi_1(1)\phi_2(1)$ for $1 \leqq u$ since the $\phi_j(u)$ have these properties. It follows that $\phi_3(u)$ is bounded and consequently of bounded variation in $-\infty < u < +\infty$. The continuity of $\phi_1(u)$ and $\phi_2(u)$ at $u = 0$ implies that of $\phi_3(u)$. It remains to discuss property (4). Here it is advantageous to express $\phi_3(u)$ by the Lebesgue integral

$$\phi_3(u) = \int_0^{\phi_2(1)} \phi_1\left(\frac{u}{v(x)}\right) dx,$$

where $v = v(x)$ is the Lebesgue inverse of the monotone function $x = \phi_2(v)$ ([7], pp. 259–260). This integral can be written

$$\phi_3(u) = \phi_1(1)\phi_2(u + 0) + \int_{\phi_2(u+0)}^{\phi_2(1)} \phi_1\left(\frac{u}{v(x)}\right) dx.$$

Replacing u by $u \pm h$ and letting h tend to zero, we get, by bounded convergence, expressions which can be written

$$\phi_3(u + 0) = \phi_1(1)\phi_2(u + 0) + \int_{\phi_2(u+0)}^{\phi_2(1)} \phi_1\left(\frac{u}{v(x)} + 0\right) dx,$$

$$\phi_3(u - 0) = [\phi_1(1) - \phi_1(1 - 0)]\phi_2(u - 0) + \phi_1(1 - 0)\phi_2(u + 0) + \int_{\phi_2(u+0)}^{\phi_2(1)} \phi_1\left(\frac{u}{v(x)} - 0\right) dx.$$

Observing that $\phi_j(u)$ $(j = 1, 2)$ are regular in $0 < u < 1$, we see that $\phi_3(u)$ will be regular at every point u which is not a point of discontinuity of $\phi_2(u)$ and will not be regular at the latter points unless $\phi_1(1 - 0) = \phi_1(1)$. This completes the proof of the lemma.

LEMMA 3.2. *With the notation and assumptions of Lemma* 3.1 *we have for every* ξ $(0 < \xi < 1)$, $\eta > 0$

$$\lim_{\eta\to 0} \mathfrak{R}^* \left\{\frac{i}{\pi} \int_{i\eta}^{\xi+i\eta} \int_0^1 \frac{d\phi_1(u)}{z - u}\, dz\right\} = \phi_1(\xi), \tag{3.4}$$

$$\lim_{\eta\to 0} \mathfrak{R}^* \left\{\frac{i}{\pi} \int_{i\eta}^{\xi+i\eta} \int_0^1 \int_0^1 \frac{d\phi_1(u)}{z - uv}\, d\phi_2(v)\, dz\right\} = \tfrac{1}{2}[\phi_3(\xi + 0) + \phi_3(\xi - 0)]. \tag{3.5}$$

Here $z = x + i\eta$ *and the asterisk indicates that in taking the real part the integrators* $\phi_j(u)$ *shall be treated as if they were real.*

Formula (3.4) is of course well known. Formula (3.5) can be proved as follows. Proceeding as in the proof of formula (2.5) we get

$$\int_0^1 \frac{d\phi_1(u)}{z - uv} = v \int_{-\infty}^{\infty} \frac{\phi_1(u)}{(z - uv)^2}\, du = \int_{-\infty}^{\infty} \frac{\phi_1(w/v)}{(z - w)^2}\, dw,$$

so that

$$\int_0^1 \left[\int_0^1 \frac{d\phi_1(u)}{z - uv}\right] d\phi_2(v) = \int_{-\infty}^{\infty} \frac{\phi_3(w)}{(z - w)^2}\, dw,$$

the interchange of the order of integration being permitted by absolute convergence of the double integral. Here we can integrate with respect to z under the sign of integration, again by absolute convergence. Taking the formal real part, we see that the left member of (3.5) becomes

$$\lim_{\eta\to 0} \frac{\eta}{\pi} \int_{-\infty}^{\infty} \left\{\frac{1}{(w - \xi)^2 + \eta^2} - \frac{1}{w^2 + \eta^2}\right\} \phi_3(w)\, dw.$$

This is the difference of two Poisson integrals for the upper half-plane and, $\phi_3(u)$ being of bounded variation, the well-known limit equals

$$\tfrac{1}{2}[\phi_3(\xi + 0) + \phi_3(\xi - 0)] - \tfrac{1}{2}[\phi_3(+0) + \phi_3(-0)].$$

The last term being zero, formula (3.5) is proved.

The proof of Theorem 3.1 is now immediate. Suppose that $\{a_n\}$ is divisible by $\{b_n\}$ and let us prove formula (3.2) for instance. Formula (2.1) is now valid. Replacing t by $-1/z$ and dividing by z in this formula, we obtain

$$\int_0^1 \frac{d\phi_a(u)}{z - u} = \int_0^1 \left[\int_0^1 \frac{d\phi_b(u)}{z - uv}\right] d\phi_c(v). \tag{3.6}$$

Putting

$$\phi(u) = \int_0^1 \phi_b(u/v)\, d\phi_c(v),$$

and applying Lemma 3.2 to formula (3.6), we get the equality

$$\phi_a(\xi) = \tfrac{1}{2}[\phi(\xi + 0) + \phi(\xi - 0)] \tag{3.7}$$

for $0 < \xi < 1$. This is equivalent to (3.2) and the exceptional set is read off from Lemma 3.1. Using (2.4) instead and proceeding in the same manner, we get (3.1).

Conversely, suppose that (3.1) holds. Assuming t in the domain T, multiplying both sides of (3.1) by $t\,du/(1 + tu)^2$ and integrating with respect to u from $-\infty$ to $+\infty$, which is clearly permitted, we get

$$\begin{aligned} t\int_{-\infty}^{\infty} \frac{\phi_a(u)}{(1 + tu)^2}\, du &= t\int_{-\infty}^{\infty} \frac{du}{(1 + tu)^2} \int_0^1 \phi_c(u/v)\, d\phi_b(v) \\ &= \int_0^1 tv \left[\int_{-\infty}^{\infty} \frac{\phi_c(w)\, dw}{(1 + tvw)^2}\right] d\phi_b(v), \end{aligned}$$

the interchange of the order of integration being allowed by the theorem of Fubini. This relation, however, reduces to formula (2.4) by virtue of (2.5). Similarly (3.2) gives (2.1). This completes the proof of Theorem 3.1.

4. **Formulation of the inclusion problem in terms of moment functions; relation to the integral equation of R. Schmidt.** Hille and Tamarkin [4] also formulated the inclusion problem in terms of moment functions, as follows.

THEOREM 4.1. *If $\{a_n\}$ and $\{b_n\}$ are regular sequences with corresponding regular moment functions $a(z)$ and $b(z)$, then $\{a_n\}$ is divisible by $\{b_n\}$ if and only if there exists a regular moment function $c(z)$ such that*

$$a(z) \equiv b(z)c(z) \quad \text{if} \quad \Re(z) \geqq 0. \tag{4.1}$$

The sufficiency of the condition is obvious since (4.1) reduces to (2.2) when $z = n$ $(n = 0, 1, 2, \cdots)$.

Conversely, let (2.2) hold and consider the function $f(z) = a(z) - b(z)c(z)$. This function vanishes for $z = n$ $(n = 0, 1, 2, \cdots)$; and if $\Re(z) \geqq 0$, then $|f(z)| < M$, where M is a constant independent of z. Now a function which is bounded for $\Re(z) \geqq 0$, is holomorphic in the interior of this half-plane, and which vanishes at the integral points, vanishes identically (Carlson [1], Theorem C; Hardy [3]). Hence (4.1) holds.

It is of interest to derive Theorem 3.1 from Theorem 4.1, and to obtain in this way the formulation of the inclusion problem in terms of the integral equation of R. Schmidt by a method which in the main is different from that used in §3.

We begin by recalling a multiplication theorem for Laplace-Stieltjes integrals.[3] Put

$$f(z) = \int_0^\infty e^{-zu}\,dA(u), \qquad g(z) = \int_0^\infty e^{-zv}\,dB(v),$$

where $A(u)$, $B(v)$ are of bounded variation on the interval $[0, \infty]$. It is no restriction to assume $A(0) = B(0) = 0$ and to suppose that for $0 < u < \infty$, $0 < v < \infty$ these functions are regular, i.e.,

$$A(u) = \tfrac{1}{2}[A(u+0) + A(u-0)], \qquad B(v) = \tfrac{1}{2}[B(v+0) + B(v-0)].$$

To simplify the proof, let us extend the definitions of $A(u)$ and $B(v)$ to negative values of u and v by setting $A(u) \equiv 0$, $B(v) \equiv 0$ for $u < 0$, $v < 0$.

An integration by parts in the first integral gives for $\Re(z) > 0$

$$f(z) = z\int_0^\infty e^{-zu} A(u)\,du.$$

Here we can of course replace the lower limit by $-\infty$. We do the same in the original integral for $g(z)$. Then for $\Re(z) > 0$

$$\begin{aligned} f(z)\,g(z) &= z\int_{-\infty}^\infty e^{-zu} A(u)\,du \cdot \int_{-\infty}^\infty e^{-zv}\,dB(v) \\ &= z\int_{-\infty}^\infty \int_{-\infty}^\infty e^{-z(u+v)} A(u)\,du\,dB(v). \end{aligned}$$

This is an absolutely convergent double integral in which we may make the change of variable $t = u + v$, $v = v$, obtaining

$$f(z)\,g(z) = z\int_{-\infty}^\infty e^{-zt}\left\{\int_{-\infty}^\infty A(t-v)\,dB(v)\right\}dt = z\int_{-\infty}^\infty e^{-zt} C(t)\,dt,$$

where

$$C(t) = \int_{-\infty}^\infty A(t-v)\,dB(v) = \int_0^t A(t-v)\,dB(v),$$

[3] Multiplication theorems for Laplace and Stieltjes-Laplace integrals abound in the literature. This particular theorem and even more general ones were stated without proof by Hille and Tamarkin in [6]. It is well known to most workers in the field. A different representation was proved by D. V. Widder ([13], p. 715).

since $A(t - v) = 0$ for $v \geqq t$, and $B(v) = 0$ for $v \leqq 0$. The properties of $C(t)$ can be read off from Lemma 3.1. Thus $C(t)$ is defined for all values of t if the integral is taken in the Lebesgue-Stieltjes sense, $C(t) = 0$ for $t \leqq 0$ and is of bounded variation on $[-\infty, +\infty]$. Further, $C(t)$ is regular either for all $t > 0$ or for all t with the exception of the points of discontinuity of $B(t)$ according as $A(+0) = 0$ or not. Let us put

$$C^*(0) = 0, \qquad C^*(t) = \tfrac{1}{2}[C(t + 0) + C(t - 0)] \qquad \text{for } t > 0.$$

We have then

$$f(z)\, g(z) = z \int_0^\infty e^{-zt} C^*(t)\, dt = \int_0^\infty e^{-zt}\, dC^*(t),$$

by another integration by parts, the second representation being valid for $\Re(z) \geqq 0$.

The function

$$C_0(t) = \int_0^t B(t - v)\, dA(v)$$

evidently has properties analogous to those of $C(t)$. In particular, $C_0(t)$ is regular either for all $t > 0$ or for all t with the exception of the points of discontinuity of $A(t)$ according as $B(+0) = 0$ or not. An integration by parts shows that

$$C_0(t) = C(t)$$

at all points where both sides are regular (see S. Saks [11], p. 102).

To prove the equivalence of (i), (ii), §1, suppose first that (i) holds, i.e., that

$$\int_0^1 u^z\, d\phi_a(u) = \int_0^1 u^z\, d\phi_b(u) \cdot \int_0^1 u^z\, d\phi_c(u).$$

Put $u = e^{-t}$, $\phi_a(u) = 1 - A(t)$, $\phi_b(u) = 1 - B(t)$, $\phi_c(u) = 1 - C(t)$ and the integrals become Laplace-Stieltjes integrals to which the above multiplication theorem applies. We then have

$$A(t) = \int_0^t B(t - s)\, dC(s),$$

the values of t which are points of discontinuity of $C(t)$ being excepted unless $B(+0) = 0$. Supposing $u = e^{-t}$, $v = e^{-s}$, we get

$$\phi_a(u) = \phi_c(u) + \int_u^1 \phi_b(u/v)\, d\phi_c(v). \tag{4.2}$$

This relation is equivalent to (3.2) and has the same exceptional set. The converse is obvious inasmuch as (4.2) implies (4.1).

5. **Hausdorff methods which include Cesàro summability of integral order.** Hille and Tamarkin [4] obtain necessary and sufficient conditions in order that $[H, \phi_a(u)] \supset (C, \alpha)$, $\alpha > 0$, by substituting in (4.2) the regular mass function for $(H, \alpha) = (C, \alpha)$:

$$\phi_b(u) = \frac{1}{\Gamma(\alpha)} \int_0^u \left\{\log \frac{1}{t}\right\}^{\alpha-1} dt, \qquad \alpha > 0.$$

Since this function is continuous, we may integrate by parts in (4.2); and if we put $u = e^{-s}$, $\phi_x(u) = 1 - p_x(s)$, the equation takes the form

$$e^s p_a(s) = \frac{1}{\Gamma(\alpha)} \int_0^s (s - t)^{\alpha-1} e^t p_c(t)\, dt. \tag{5.1}$$

Consequently, $[H, \phi_a(u)] \supset (C, \alpha)$ if and only if (5.1) has a solution $p_c(s)$ which satisfies the conditions:

(i) $p_c(s)$ is of bounded variation on the interval $(0, \infty)$,

(ii) $\lim_{s \to \infty} p_c(s) = 1$.

To these should be added the requirement $p_c(0) = 0$, which, however, is not essential since the value of the right member of (5.1) cannot be affected by changing the value of $p_c(s)$ at $s = 0$.

In case α is an integer, $\alpha = n \geqq 1$, one may obtain conditions on $p_a(s)$ which are necessary and sufficient for $[H, \phi_a(u)] \supset (C, n)$. In the statement and proof of this theorem we simplify the notation by replacing $p_a(s)$ and $p_c(s)$ by $p(s)$ and $\pi(s)$ respectively.

THEOREM 5.1. *Necessary and sufficient conditions that* $[H, \phi(u)] \supset (C, n)$, *where n is an integer* $\geqq 1$, *are*

(i) $p(s)$ *is absolutely continuous and has absolutely continuous derivatives of orders* $\leqq n - 1$ *for* $0 < s < \infty$,

(ii) $p^{(n-1)}(s)$ *has a finite right-hand derivative* $p_r^{(n)}(s)$ *for* $s \geqq 0$, *and a finite left-hand derivative* $p_l^{(n)}(s)$ *for* $s > 0$,

(iii) $p(s), p'(s), \cdots, p^{(n-1)}(s), p_r^{(n)}(s), p_l^{(n)}(s)$ *are of bounded variation in the interval* $(0, \infty)$,

(iv) $p(s), p'(s), \cdots, p^{(n-1)}(s)$ *tend to zero as* $s \to 0$,

(v) $1 - p(s), p'(s), \cdots, p^{(n-1)}(s), p_r^{(n)}(s), p_l^{(n)}(s)$ *tend to zero as* $s \to \infty$.

We rewrite Theorem 5.1 translating the conditions on $p(s)$ into conditions on $\phi(u)$.[4]

THEOREM 5.2. *Necessary and sufficient conditions that* $[H, \phi(u)] \supset (C, n)$, *where n is an integer* $\geqq 1$, *are*

[4] The wording of the corresponding theorem in Hille and Tamarkin [4], p. 576, is partly incorrect. The subsequent proof is essentially that originally intended by those authors, who have permitted us to publish it here. We have tried to state the conditions as concisely as possible, but some parts of the conditions may be redundant.

(i) *$\phi(u)$ is absolutely continuous and has absolutely continuous derivatives of orders $\leqq n - 1$ for $0 < u \leqq 1$,*

(ii) *$\phi^{(n-1)}(u)$ has a finite right-hand derivative $\phi_r^{(n)}(u)$ for $0 < u < 1$, and a finite left-hand derivative $\phi_l^{(n)}(u)$ for $0 < u \leqq 1$.*

(iii) *$\phi(u), u\phi'(u), \cdots, u^{n-1}\phi^{(n-1)}(u), u^n\phi_r^{(n)}(u), u^n\phi_l^{(n)}(u)$ are of bounded variation in the interval $(0, 1)$,*

(iv) *$1 - \phi(u), \phi'(u), \cdots, \phi^{(n-1)}(u)$ tend to zero as $u \to 1$,*

(v) *$\phi(u), u\phi'(u), \cdots, u^{n-1}\phi^{(n-1)}(u), u^n\phi_r^{(n)}(u), u^n\phi_l^{(n)}(u)$ tend to zero as $u \to 0$.*

The proof of Theorem 5.1 is based on an induction argument. Accordingly we shall prove the theorem first for the case $n = 1$. We have to consider the equation

$$e^s p(s) = \int_0^s e^t \pi(t)\,dt.$$

We assume that this equation has a solution $\pi(s)$ of bounded variation on the interval $(0, \infty)$ such that $\pi(s) \to 1$ as $s \to \infty$. It is at once clear that $p(s)$ must be absolutely continuous and that $p(s) \to 0$ as $s \to 0$. Thus conditions (i) and (iv) are satisfied.

Suppose $h > 0$. Then, for $s \geqq 0$, we have

$$\frac{1}{h}[e^{s+h}p(s+h) - e^s p(s)] = \frac{1}{h}\int_s^{s+h} e^t \pi(t)\,dt \to e^s \pi(s+0) \tag{5.2}$$

as $h \to 0$. This statement is valid since $\pi(s)$ is of bounded variation and hence has a right-hand limit everywhere. Moreover, we have

$$\frac{1}{h}[e^{s+h}p(s+h) - e^s p(s)] = \frac{1}{h}(e^{s+h} - e^s)p(s+h) + \frac{e^s}{h}[p(s+h) - p(s)].$$

Here, the first term on the right tends to $e^s p(s)$ since $p(s)$ is continuous, and from (5.2) the limit of the term on the left exists. It follows that the second term on the right tends to the limit $e^s p_r'(s)$ as $h \to 0$. This implies the existence of $p_r'(s)$ for $s \geqq 0$. Indeed, from (5.2) we have $p(s) + p_r'(s) = \pi(s + 0)$, $s \geqq 0$. In a similar fashion we may prove the existence of a unique finite left-hand derivative $p_l'(s)$ and derive the relation $p(s) + p_l'(s) = \pi(s - 0)$, $s > 0$. Condition (ii) has thus been fulfilled.

By assumption $\pi(s)$ and $p(s)$ are of bounded variation in the interval $(0, \infty)$. This implies that $\pi(s + 0)$ and $\pi(s - 0)$ are of bounded variation in the interval $(0, \infty)$. Since $p_r'(s) = \pi(s + 0) - p(s)$, $p_l'(s) = \pi(s - 0) - p(s)$, we see that condition (iii) is satisfied. It remains to prove that condition (v) is satisfied.

It follows by assumption that $1 - p(s) \to 0$ as $s \to \infty$, since $\phi(u)$ defines a regular method of summation. We have also by assumption $\pi(s) \to 1$ as $s \to \infty$. Hence, using the relations derived above, we have $p_r'(s) \to 0$ and $p_l'(s) \to 0$ as $s \to \infty$. Thus condition (v) is fulfilled, and the necessity of the conditions has been established for the case $n = 1$.

To prove the sufficiency for the case $n = 1$, let $p(s)$ satisfy the conditions of the theorem, and define $\pi(s)$ by the equation

$$\pi(s) = p(s) + \tfrac{1}{2}[p_r'(s) + p_l'(s)].$$

This function evidently satisfies the requirements

$$(5.3)\qquad \begin{aligned}&\text{(i) } \pi(s) \text{ is of bounded variation in } (0, \infty),\\ &\text{(ii) } \lim_{s\to\infty} \pi(s) = 1.\end{aligned}$$

It remains to be shown that this function satisfies the integral equation

$$\int_0^s e^t \pi(t)\, dt = e^s p(s).$$

We have ([9], pp. 507–510)

$$\begin{aligned}\pi(s + 0) &= p(s + 0) + \tfrac{1}{2}[p_r'(s + 0) + p_l'(s + 0)]\\ &= p(s) + p_r'(s),\end{aligned}$$

and therefore

$$\begin{aligned}\int_0^s e^t \pi(t)\, dt &= \int_0^s e^t \pi(t + 0)\, dt\\ &= \int_0^s e^t [p(t) + p_r'(t)]\, dt = e^s p(s),\end{aligned}$$

where we have set $p(0) = p(+0)$.

To complete the induction argument let us assume that the theorem has been proved for $n \leqq k$, and then prove it for $n = k + 1$. Suppose then that $[H, \phi(u)] \supset (C, k + 1)$. Since $(C, k + 1) \supset (C, k)$, the function $p(s)$ must satisfy the conditions of the theorem for $n \leqq k$. We have further

$$e^s p(s) = \frac{1}{\Gamma(k + 1)} \int_0^s (s - t)^k e^t \pi(t)\, dt.$$

After $k - 1$ differentiations we get

$$e^s[p(s) + C_{k-1,1}p'(s) + \cdots + p^{(k-1)}(s)] = \int_0^s (s - t)e^t \pi(t)\, dt.$$

Taking a right-hand derivative we get

$$e^s[p(s) + C_{k,1}p'(s) + \cdots + C_{k,k-1}p^{(k-1)}(s) + p_r^{(k)}(s)] = \int_0^s e^t \pi(t)\, dt,$$

and the same expression holds if we replace r by l. Hence $p_r^{(k)}(s) = p_l^{(k)}(s) = p^{(k)}(s)$, an absolutely continuous function. Clearly $p^{(k)}(s) \to 0$ as $s \to 0$. Since $\pi(s)$ has left- and right-hand limits, we prove the existence of $p_r^{(k+1)}(s)$ and

$p_l^{(k+1)}(s)$, the former for $s \geqq 0$, the latter for $s > 0$, by the same method as when $n = 1$. Hence

$$p(s) + C_{k+1,1}p'(s) + \cdots + C_{k+1,k}p^{(k)}(s) + p_r^{(k+1)}(s) = \pi(s + 0),$$

$$p(s) + C_{k+1,1}p'(s) + \cdots + C_{k+1,k}p^{(k)}(s) + p_l^{(k+1)}(s) = \pi(s - 0).$$

Since $\pi(s + 0)$ and $\pi(s - 0)$ are of bounded variation on $(0, \infty)$, the same is true of $p_r^{(k+1)}(s)$ and $p_l^{(k+1)}(s)$. Moreover, since $\pi(s) \to 1$ as $s \to \infty$, we conclude that also $p_r^{(k+1)}(s)$ and $p_l^{(k+1)}(s)$ tend to zero as $s \to \infty$. This completes the proof of the necessity of the conditions for $n = k + 1$.

To prove the sufficiency we proceed in the same way as for the case $n = 1$. Given a $p(s)$ satisfying the conditions of the theorem for $n \leqq k$, we define $\pi(s)$ by the equation

$$\pi(s) = p(s) + C_{k+1,1}p'(s) + \cdots + C_{k+1,k}p^{(k)}(s) + \tfrac{1}{2}[p_r^{(k+1)}(s) + p_l^{(k+1)}(s)].$$

This function satisfies conditions (5.3). Also,

$$\int_0^s e^t\pi(t)\,dt = \int_0^s e^t[p(t) + C_{k+1,1}p'(t) + \cdots + C_{k+1,k}p^{(k)}(t) + p_r^{(k+1)}(t)]\,dt$$
$$= e^s p(s).$$

This completes the proof of the theorem.

6. **Inclusion problems in the difference matrix.** If $\{x_n\}$ is a regular sequence, other than the unit sequence $(1, 1, 1, \cdots)$, then the rows in the *difference matrix* [2]

$$(\Delta^m x_n) = \begin{pmatrix} x_0 & x_1 & x_2 & \cdots \\ \Delta x_0 & \Delta x_1 & \Delta x_2 & \cdots \\ \Delta^2 x_0 & \Delta^2 x_1 & \Delta^2 x_2 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

are all essentially regular, that is, they can be made regular by dividing all terms by the first terms. The columns and diagonals are also essentially regular if the mass function for the base sequence $\{x_n\}$ is continuous at $u = 1$. Thus, from the difference matrix generated by a regular base sequence there arises an infinitude of Hausdorff methods of summability, between any two of which the problem of relative inclusion immediately presents itself. These problems have been discussed elsewhere for several special cases [2]. We mention here one fact which lends interest to the problem, namely: in some cases the Hausdorff methods defined by the rows *increase* in efficiency with the depth of the row in the matrix, while in other cases methods defined by two consecutive rows are not comparable, that is, neither includes the other.

Another kind of inclusion problem arises from the simultaneous consideration of *two* difference matrices. If the Hausdorff methods defined by the base se-

quences of the two matrices are equivalent, what can be said about the methods defined by other pairs of corresponding sequences?

We shall simplify the statement of the next and later theorems by agreeing that the words "the sequence $\{a_n\}$ includes the sequence $\{b_n\}$, or is equivalent to the sequence $\{b_n\}$" shall mean that the Hausdorff method defined by the first sequence includes, or is equivalent to, the Hausdorff method defined by the second. We may also speak of one row of a difference matrix including or being equivalent to another.

THEOREM 6.1. *Given two difference matrices $(\Delta^m a_n)$ and $(\Delta^m b_n)$ in which the row sequences are all essentially regular, and such that the two base sequences are equivalent, any one of the three following possibilities may actually occur:*

(i) *the second rows of the two matrices are equivalent;*

(ii) *the second row of the first matrix includes the second row of the second, but not vice versa;*

(iii) *the second rows of the two matrices are not comparable, that is, neither includes the other.*

That (i) is possible will follow from the work in the succeeding sections.

To show that (ii) may be realized, take the base sequences to be $\{(200n + 1)/(n + 1)^2\}$ and $\{1/(n + 1)\}$. It is easy to see that these are equivalent sequences. The sequences in the second rows of the matrices are $\{(200n^2 + 202n - 197)/(n + 1)^2(n + 2)^2\}$ and $\{1/(n + 1)(n + 2)\}$. The first of these sequences is readily seen to be divisible by the second; but the quotient of the second by the first is $\{(n + 1)(n + 2)/(200n^2 + 202n - 197)\}$. In order for the second sequence to be divisible by the first the function $(z + 1)(z + 2)/(200z^2 + 202z - 197)$ must be a regular moment function. This is impossible inasmuch as this function has a pole in the right half-plane at the point $z = .6+$.

An example showing that (iii) may be realized can be constructed along these same lines.

7. **Proof of the relation $(H_m, \alpha) \supset (C_m, \alpha)$, where α is real and positive, by means of the method of integral equations.** The regular sequences defining (H, α) and (C, α) are

$$\{(n + 1)^{-\alpha}\}, \qquad \{1/C_{n+\alpha,n}\} \tag{7.1}$$

respectively. It will be convenient to denote by (H_m, α) and (C_m, α) the methods of summation defined by the sequences in the m-th rows of the difference matrices for the methods (H, α) and (C, α), respectively, i.e., the matrices having (7.1) as base sequences. We shall employ the integral equation (4.2) to prove the theorem which follows.

THEOREM 7.1. *If α is real and positive, then $(H_m, \alpha) \supset (C_m, \alpha)$ for $m = 2, 3, 4, \cdots$.*

Inasmuch as $(C_m, \alpha) = (C, \alpha + m - 1) \approx (H, \alpha + m - 1)$, this is the same as proving that

$$(H_m, \alpha) \supset (H, \alpha + m - 1) \qquad (m = 2, 3, 4, \cdots). \tag{7.2}$$

Since the mass function for (H, α) is

$$\frac{1}{\Gamma(\alpha)} \int_0^u \left(\log \frac{1}{t}\right)^{\alpha-1} dt,$$

it follows that the mass function for (H_m, α) is

$$\phi_a(u) = A \int_0^u (1 - t)^{m-1} \left(\log \frac{1}{t}\right)^{\alpha-1} dt, \tag{7.3}$$

where A is a constant so determined that $\phi_a(1) = 1$. Put

$$f(s) = e^s(1 - \phi_a(u)), \qquad u = e^{-s},$$

and the integral equation for the relation (7.3) is, by (4.2):

$$f(s) = \frac{1}{\Gamma(\mu + \beta - 1)} \int_0^s (s - t)^{\mu+\beta-2} e^t p_c(t)\, dt, \tag{7.4}$$

where μ, β are so determined that

$$\alpha + m = \mu + \beta \qquad (\mu \text{ an integer} > 2,\ 0 \leqq \beta < 1).$$

Let us assume that (7.4) has a continuous solution $p_c(t)$, and put

$$V_1(t) = \int_0^t e^x p_c(x)\, dx,\ V_2(t) = \int_0^t V_1(x)\, dx,\ \cdots,\ V_{\mu-2}(t) = \int_0^t V_{\mu-3}(x)\, dx.$$

Then, after $\mu - 2$ successive integrations by parts, the equation (7.4) takes the form

$$f(s) = \frac{1}{\Gamma(\beta + 1)} \int_0^s (s - t)^\beta V_{\mu-2}(t)\, dt \qquad (0 \leqq \beta < 1). \tag{7.5}$$

There are two cases to be considered according as $\beta = 0$ or $\beta > 0$. If $\beta = 0$, then on differentiating (7.5) with respect to s we get

$$V_{\mu-2}(s) = f(s),$$

and consequently

$$e^s p_c(s) = f^{(\mu-1)}(s).$$

Let $p_a(s) = 1 - \phi_a(u)$, $u = e^{-s}$. Then we find that

$$p_c(s) = p_a(s) + C_{\mu-1,1} p_a'(s) + C_{\mu-1,2} p_a''(s) + \cdots + C_{\mu-1,\mu-1} p_a^{(\mu-1)}(s).$$

One may verify that this function $p_c(s)$ is a continuous function of s which satisfies the integral equation (7.4). Moreover, $\lim_{s\to\infty} p_c(s) = 1$, and $p_c(s)$ is of bounded variation on $(0, \infty)$. Therefore $(H_m, \alpha) \supset (H, \alpha + m - 1)$ in this case.

If $0 < \beta < 1$, we put

$$V_{\mu-1}(t) = \int_0^t V_{\mu-2}(x)\,dx,$$

and integrate by parts once more in (7.5), getting

$$f(s) = \frac{1}{\Gamma(\beta)} \int_0^s \frac{V_{\mu-1}(t)\,dt}{(s-t)^{1-\beta}},$$

which is an Abel integral equation. The solution is

$$V_{\mu-1}(z) = \frac{1}{\pi}\Gamma(\beta)\sin\pi\beta\,\frac{d}{dz}\int_0^z \frac{f(s)\,ds}{(z-s)^\beta} = \frac{1}{\Gamma(1-\beta)}\int_0^z \frac{f'(s)\,ds}{(z-s)^\beta}.$$

On differentiating this $\mu - 1$ times with respect to z, we then have

$$p_c(z) = \frac{1}{\Gamma(1-\beta)}\,e^{-z}\int_0^z \frac{f^{(\mu)}(s)\,ds}{(z-s)^\beta}.$$

It is not difficult to see that $p_c(z)$ is continuous and satisfies the integral equation, and that $p_c(z)$ is of bounded variation on $(0, \infty)$.

To show that $\lim_{z\to\infty} p_c(z) = 1$, we first write $p_c(z)$ in the form

$$p_c(z) = \frac{1}{\Gamma(1-\beta)}\,e^{-z}\left\{\int_0^z \frac{e^s ds}{(z-s)^\beta} - \int_0^z \frac{e^s\phi_\alpha(e^{-s})\,ds}{(z-s)^\beta}\right.$$
$$\left. + \int_0^z e^s[C_{\mu,1}p_\alpha'(s) + C_{\mu,2}p_\alpha''(s) + \cdots + C_{\mu,\mu}p_\alpha^{(\mu)}(s)]\,\frac{ds}{(z-s)^\beta}\right\}.$$

Designate the integrals in this expression, exclusive of constant multipliers, by I_1, I_2, I_3. Then

$$I_1 = e^{-z}\int_0^z \frac{e^s\,ds}{(z-s)^\beta} = \int_0^z \frac{e^{s-z}ds}{(z-s)^\beta} = \int_0^z e^{-w}w^{-\beta}\,dw,$$

and

$$\lim_{z\to\infty}\frac{1}{\Gamma(1-\beta)}I_1 = 1.$$

It remains to prove that I_2 and I_3 tend to zero as $z \to \infty$. We have

$$\phi_\alpha(e^{-s}) = A\int_0^{e^{-s}} (1-t)^{m-1}\left(\log\frac{1}{t}\right)^{\alpha-1} dt$$
$$= A\int_s^\infty (1-e^{-x})^{m-1}x^{\alpha-1}e^{-x}\,dx$$
$$< A\int_s^\infty x^{\alpha-1}e^{-x}\,dx < C(\alpha)s^{\alpha-1}e^{-s},$$

if $s \geqq 1$, while $\phi_a(e^{-s}) < C(\alpha)$ for all s.[5] Then for $z \geqq 1$

$$I_2 < C(\alpha)e^{-z}\left\{\int_0^z \frac{s^{\alpha-1}\,ds}{(z-s)^\beta} + \int_0^1 \frac{ds}{(z-s)^\beta}\right\}$$

$$< C(\alpha)e^{-z}\left[\mathrm{B}(\alpha, 1-\beta)z^{\alpha-\beta} + \frac{1}{1-\beta}z^{1-\beta}\right].$$

Accordingly, $\lim_{z\to\infty} I_2 = 0$.

Next we consider I_3. It will suffice to show that an arbitrary term of I_3,

$$J_i = C_{\mu,i}e^{-z}\int_0^z \frac{e^s p_a^{(i)}(s)}{(z-s)^\beta}\,ds \qquad (i = 1, 2, 3, \cdots, \mu),$$

tends to zero as $z \to \infty$. If $p_a^{(i)}(s)$ is computed in terms of derivatives of its component parts, we obtain a linear combination of terms of the form $e^{-js}(1 - e^{-s})^{m-j}s^{\alpha-k-1}$, $j + k \leqq i$, $1 \leqq j$, every one of which is dominated for all $s > 0$ by an expression of the form $e^{-s}s^\gamma$ with $-1 < \beta - 1 \leqq \gamma \leqq \mu + \beta - 2$. The term with the largest value of γ gives the largest contribution to the integral of the dominant for large values of z. Hence

$$|J_i| < C(\alpha, m)e^{-z}\int_0^z \frac{s^{\mu+\beta-2}\,ds}{(z-s)^\beta} = C(\alpha, m)\mathrm{B}(\mu + \beta - 1, 1 - \beta)e^{-z}z^{\mu-1}.$$

Accordingly, $\lim_{z\to\infty} J_i = 0$, and thus $\lim_{z\to\infty} I_3 = 0$.

Collecting our information on I_1, I_2, I_3 we conclude that $\lim_{z\to\infty} p_c(z) = 1$. Since the last of the regularity requirements on $p_c(t)$ has been met, our proof is completed.

We note that the method of proof used in this section, unlike the method used in the following section, enables us to exhibit the mass function for the quotient of the moment sequences involved. For example, in the case $\alpha = 2$, $m = 2$ we have

$$\phi_c(u) = \begin{cases} \frac{1}{3}u^2(2\log u + 5), & 0 \leqq u \leqq 1, \\ 1, & u = 1. \end{cases}$$

8. **Proof of the relation $(H_m, \alpha) \supset (C_m, \alpha)$ for the real part of α positive by the method of moment functions.** The regular moment functions for (H_m, α) and $(H, \alpha + m - 1) \approx (C_m, \alpha)$ are $\mu_1(z)$ and $\mu_2(z)$, where

$$A\mu_1(z) = (z+1)^{-\alpha} - C_{m-1,1}(z+2)^{-\alpha} + \cdots + (-1)^{m-1}(z+m)^{-\alpha},$$

$$\mu_2(z) = (z+1)^{-\alpha-m+1},$$

and A is a constant chosen so that $\mu_1(0) = 1$. It is required to show that $\mu_1(z)/\mu_2(z)$ is a regular moment function.

[5] We use the symbol $C(\alpha)$ to denote an unspecified positive constant, not always the same, depending upon the particular value of α. Similarly for $C(\alpha, m)$ below.

Let

$$\mu(z) = \frac{\mu_1(z)}{\mu_2(z)} - A_0,$$

where A_0 is a constant such that $\mu(z)$ vanishes at $z = \infty$. It will suffice to prove that $\mu(z)$ is a moment function

$$\int_0^1 u^z \, d\phi(u), \tag{8.1}$$

where $\phi(u)$ is of bounded variation on (0, 1) and is continuous at $u = 0$.

It is clear first that the function $\mu(z)$ is holomorphic for $\Re(z) \geqq -1 + \delta$, $\delta > 0$. It is also holomorphic for $|z| > m$, and since it vanishes for $z = \infty$, we have

$$\mu(z) = \frac{A_1}{z} + \frac{A_2}{z^2} + \frac{A_3}{z^3} + \cdots, \qquad |z| > m.$$

From this it follows that

$$|z + 1| \cdot |\mu(z)| \leqq C \quad \text{for} \quad |z| \geqq m + 1.$$

Also, since $\mu(z)$ is holomorphic for $\Re(z) \geqq -1 + \delta$, $|z| \leqq m + 1$, we conclude that

$$|\mu(z)| \leqq \frac{C}{|z+1|} = \frac{C}{[(x+1)^2 + y^2]^{\frac{1}{2}}}, \qquad z = x + iy, \quad x \geqq -1 + \delta,$$

and therefore

$$\int_{-\infty}^{\infty} |\mu(x + iy)|^2 \, dy \leqq C^2 \int_{-\infty}^{\infty} \frac{dy}{(x+1)^2 + y^2} = \frac{C^2\pi}{x+1} \leqq \frac{C^2\pi}{\delta}, \qquad x \geqq -1 + \delta.$$

Now[6] if $f(z)$ is holomorphic for $x > a$ and $\int_{-\infty}^{\infty} |f(x + iy)|^2 \, dy \leqq M$ for $x \geqq a$, then

$$f(z) = \int_0^{\infty} e^{-zt} F(t) \, dt,$$

where

$$\frac{1}{2\pi} \int_0^{\infty} e^{-2at} |F(t)|^2 \, dt = \int_{-\infty}^{\infty} |f(a + iy)|^2 \, dy.$$

[6] The identity of the class $H_2(0)$ with the class of Laplace transforms of quadratically integrable density functions appears to have been explicitly stated for the first time by Paley and Wiener ([8], pp. 8–9), though they had used the facts also in earlier communications. For the present proof we need merely the fact that if $f(z) \in H_2(0)$, then $f(z + \epsilon)$ is a regular moment function for every $\epsilon > 0$. This fact was stated by Hille and Tamarkin ([5], p. 903), for $f(z) \in H_p(0)$, $1 \leqq p \leqq 2$, with some indication of a proof.

Applying this to the present case, taking $a = -\beta$ $(0 < \beta < 1)$, we have

$$\mu(z) = \int_0^\infty e^{-zt} F(t)\, dt,$$

and

$$\int_0^\infty e^{2\beta t} \mid F(t) \mid^2 dt$$

converges. But

$$\left(\int_0^w \mid F(t) \mid dt\right)^2 \leqq \int_0^w e^{-2\beta t}\, dt \cdot \int_0^w e^{2\beta t} \mid F(t) \mid^2 dt,$$

whence it follows that $\int_0^\infty \mid F(t) \mid dt$ exists. Thus, if we put

$$p(t) = -\int_t^\infty F(s)\, ds,$$

then $p(t)$ is a continuous function of bounded variation on $(0, \infty)$, which tends to zero as $t \to \infty$, and

$$\mu(z) = \int_0^\infty e^{-zt}\, dp(t) = \int_0^1 u^z\, d\phi(u),$$

where $u = e^{-t}$, $\phi(u) = -p\,(-\log u)$. Hence $\mu(z)$ is a moment function of the form (8.1), where $\phi(u)$ is of bounded variation on (0, 1) and is continuous at $u = 0$.

9. **Discussion of the equivalence relation** $(H_m, \alpha) \approx (C_m, \alpha)$. In the preceding sections we have shown that

$$(H_m, \alpha) \supset (C_m, \alpha) \qquad (m = 1, 2, 3, \cdots ; R(\alpha) > 0).$$

The problem of establishing the opposite relation, namely,

$$(C_m, \alpha) \supset (H_m, \alpha)$$

resolves itself into the problem of showing that the function

$$\mu_1(z) = (1 + z)^{-\alpha} - C_{m-1,1}(2 + z)^{-\alpha} + \cdots + (-1)^{m-1}(m + z)^{-\alpha}$$

does not vanish for $\Re(z) \geqq 0$. This we have succeeded in doing only for the values of m and α indicated in the following table.

(9.1)

m	α
1	positive real part
2	positive real part
3	1, 2, 3, $\cdots$, 10
4	1, 2, 3, 4
5	1, 2, 3, 4
$\geqq 6$	1, 2, 3.

For these values of m and α we therefore have

$$(C_m, \alpha) \approx (H_m, \alpha) \approx (H, \alpha + m - 1).$$

For $m = 1$ this is simply the well-known equivalence relation $(H, \alpha) \approx (C, \alpha)$.

For $m = 2$, $\mu_1(z) = (1 + z)^{-\alpha} - (2 + z)^{-\alpha} \neq 0$, $\Re(z) \geqq 0$, so that $(H_2, \alpha) \subset (H, \alpha + 1)$, and therefore $(H_2, \alpha) \approx (H, \alpha + 1)$.

It will be convenient in what follows to write $\mu_1(z) = \mu_1(m, \alpha, z)$. Then

$$\mu_1(m, 1, z) = (1 + z)^{-1} - C_{m-1,1}(2 + z)^{-1} + \cdots + (-1)^{m-1}(m + z)^{-1}$$

$$= \frac{(m-1)!}{(1 + z)(2 + z) \cdots (m + z)} \neq 0,$$

and therefore $(H_m, 1) \subset (H, m)$, $(H_m, 1) \approx (H, m)$. Now

$$\mu_1(m, 2, z) = -D_z\mu_1(m, 1, z) = \frac{(m-1)!}{(1 + z)(2 + z) \cdots (m + z)} \sum_{k=1}^{m} \frac{1}{k + z}.$$

This expression has only real and negative zeros, one in each of the intervals $(-1, -2), (-2, -3), \cdots, (-m + 1, -m)$. Hence $(H_m, 2) \subset (H, m + 1)$, $(H_m, 2) \approx (H, m + 1)$. To continue this process, we have

$$\mu_1(m, 3, z) = - D_z\mu_1(m, 2, z) = \frac{(m-1)!}{(1 + z)(2+z) \cdots (m + z)} \sum_{1 \leqq r \leqq s \leqq m} \frac{1}{(r + z)(s + z)}.$$

Now

$$\sum \frac{1}{(r + z)(s + z)} = \sum \frac{\{(r + x)(s + x) - y^2\} - iy \{r + s + 2x\}}{|(r + z)(s + z)|^2},$$

where $z = x + iy$. When $y = 0$, the real part of this expression is positive if $x \geqq 0$; while if $y \neq 0$, the imaginary part is different from 0 for $x \geqq 0$. Hence $\mu_1(m, 3, z) \neq 0$ for $\Re(z) \geqq 0$, and consequently $(H_m, 3) \subset (H, m + 2)$, $(H_m, 3) \approx (H, m + 2)$.

Unfortunately, the calculations involved in proceeding along these lines become increasingly difficult, so that this method is of somewhat limited usefulness. However, by use of this method we can employ a theorem of Pólya ([10], p. 37) to show that $(H_m, \alpha) \subset (H, \alpha + m - 1)$ for each fixed m if α is sufficiently large and integral. Let K denote the set of zeros of the functions $\mu_1(m, 1, z)$, $D_z\mu_1(m, 1, z)$, $D_z^2\mu_1(m, 1, z)$, $\cdots$. Then, according to the theorem of Pólya, the set of limit points of K consists of the points of the vertical lines $\Re(z) = -\frac{3}{2}$, $\Re(z) = -\frac{5}{2}$, $\cdots$, $\Re(z) = \frac{1}{2}(1 - 2m)$. Consequently, for a fixed m, at most a finite number of the functions $\mu_1(m, 1, z)$, $\mu_1(m, 2, z)$, $\mu_1(m, 3, z)$, $\cdots$ can vanish for $\Re(z) \geqq 0$, so that $(H_m, \alpha) \subset (H, m + \alpha - 1)$ if α is sufficiently large.

The other cases in the table (9.1) have been established by direct investigation of the equation $\mu_1(z) = 0$, by means of purely algebraic methods. Such methods involve a considerable amount of computation and apparently do not lead to a general solution of the problem.

It is perhaps of some interest to set up the integral equation formulation of the problem. Let $\phi_a(u)$, $\phi_b(u)$ denote the regular mass functions for $(H, \alpha + m - 1)$ and (H_m, α), respectively. Then

$$\phi_a(u) = \frac{1}{\Gamma(\alpha + m - 1)} \int_0^u \left(\log \frac{1}{t}\right)^{\alpha+m-2} dt,$$

$$\phi_b(u) = A \int_0^u (1 - t)^{m-1} \left(\log \frac{1}{t}\right)^{\alpha-1} dt,$$

where A is a normalizing factor such that $\phi_b(1) = 1$. The integral equation for the relation $(H_m, \alpha) \subset (H, \alpha + m - 1)$ is then

$$p_a(s) = A \int_0^s (s - t)^{\alpha-1} e^{t-s}(1 - e^{t-s})^{m-1} p_c(t)\, dt,$$

where $p_x(s) = 1 - \phi_x(u)$, $u = e^{-s}$.

When $\alpha = 2$, $m = 2$, we find that

$$p_c(t) = 1 - \tfrac{1}{8}e^{-t}(6 - e^{-\frac{1}{2}t});$$

and when $\alpha = 2$, $m = 3$, the solution is

$$p_c(t) = 1 - \frac{e^{-t}}{2} + \frac{e^{-\frac{1}{2}t}}{21}\left[2 \cos \frac{3^{\frac{1}{2}}t}{6} - 3 \sin \frac{3^{\frac{1}{2}}t}{6}\right].$$

References

1. F. Carlson, *Sur une classe de séries de Taylor*, Dissertation, Uppsala, 1914, 76 pp.
2. H. L. Garabedian and H. S. Wall, *Hausdorff methods of summation and continued fractions*, Transactions of the American Mathematical Society, vol. 48(1940), pp. 185–207.
3. G. H. Hardy, *On two theorems of F. Carlson and S. Wigert*, Acta Mathematica, vol. 42(1920), pp. 327–331.
4. E. Hille and J. D. Tamarkin, *Questions of relative inclusion in the domain of Hausdorff means*, Proceedings of the National Academy of Sciences, vol. 19(1933), pp. 573–577.
5. E. Hille and J. D. Tamarkin, *On moment functions*, Proceedings of the National Academy of Sciences, vol. 19(1933), pp. 902–908.
6. E. Hille and J. D. Tamarkin, *On the theory of Laplace integrals*, II, Proceedings of the National Academy of Sciences, vol. 20(1934), pp. 140–144.
7. H. Lebesgue, *Leçons sur l'Intégration et la Recherche des Fonctions Primitives*, 2d edition, Paris, 1928.
8. R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, American Mathematical Society Colloquium Publications, vol. 19, 1934.
9. J. Pierpont, *The Theory of Functions of Real Variables*, vol. 2, New York, 1912.
10. G. Pólya, *Über die Nullstellen sukzessiver Derivierten*, Mathematische Zeitschrift, vol. 12(1922), pp. 36–60.
11. S. Saks, *Theory of the Integral*, Warszawa-Lwów, 1937.
12. R. Schmidt, *Über divergente Folgen und lineare Mittelbildungen*, Mathematische Zeitschrift, vol. 22(1925), pp. 89–152.
13. D. V. Widder, *A generalization of Dirichlet's series and of Laplace's integrals by means of a Stieltjes integral*, Transactions of the American Mathematical Society, vol. 31(1929), pp. 694–743.

Northwestern University and Yale University.

A CORRECTION

By J. Shohat

In the paper *Laguerre polynomials and the Laplace transform*, this Journal, vol. 6(1940), pp. 615–626, in the eighth line from the bottom of page 618 add "$= 0$".

rnal,
add

FUNCTIONS OF BOUNDED VARIATION IN TWO VARIABLES

By M. S. Macphail

1. In a recent paper on non-absolutely convergent integrals, R. L. Jeffery gave the definition of a class V_1 of functions on the rectangle $R:(a, b; c, d)$ $(a \leqq x \leqq c, b \leqq y \leqq d)$.[1] A function $F(x, y)$ is in class V_1 on R if there exist on R a single-valued function $f(x, y)$ and a sequence of summable functions $s_n(x, y)$ tending to $f(x, y)$, such that $\int_e s_n(x, y)\, dx\, dy$ is bounded for all values of n and all measurable sets $e \subset R$, and

$$F(x, y) = \lim_{n \to \infty} \int_{(a,b;x,y)} s_n(x, y)\, dx\, dy \tag{A}$$

for all points (x, y) on R. To save repetition, we shall take R as the fundamental rectangle throughout this paper. All points and sets mentioned will be understood to lie on R; all functions will be taken as defined over R, and all statements regarding summability, measurability, functional class, and so forth, will be understood to apply over R.

It was shown in Theorem II of J that any function in class V_1 is in the class H of functions of bounded variation in the Hardy-Krause sense. The precise relationship of V_1 to the various definitions of bounded variation that have been proposed was, however, not determined; it is part of our present purpose to show that a slightly modified form of V_1 is equivalent to the class V of Vitali, which includes H.

A further definition was given of functions in class V_2. This is entirely similar to the above, except that the condition that $\int_e s_n(x, y)\, dx\, dy$ be bounded in n and e is not imposed. It was shown in Lemma VIII and Theorem XI of J that if $F(\omega)$ is a continuous function of intervals ω for which the strong derivative $F'_s(x, y)$ is finite everywhere, then there exists a sequence of functions $s_n(x, y)$ tending to $F'_s(x, y)$ for which $\int_{(a,b;x,y)} s_n(x, y)\, dx\, dy$ tends to $F(a, b; x, y)$. In other words $\phi(x, y) = F(a, b; x, y)$ is in class V_2 with $f(x, y) = F'_s(x, y)$. The question was raised whether the strong derivative could be replaced by the ordinary derivative in case the strong derivative was not known to exist. We answer this question by proving that given $F(\omega)$ a continuous additive function

Received January 18, 1940; in revised form, January 5, 1941.

[1] R. L. Jeffery, *Functions of bounded variation and non-absolutely convergent integrals in two or more dimensions*, this Journal, vol. 5(1939), pp. 753–774. Here the point (a, b) is taken as $(0, 0)$. This paper will be referred to as J.

of intervals and $f(x, y)$ any measurable function, we can find a sequence $s_n(x, y)$ tending to $f(x, y)$ for which $\int_{(a,b;x,y)} s_n(x, y)\,dx\,dy$ tends to $F(a, b; x, y)$. In particular we may take $f(x, y)$ equal to the ordinary derivative $F'(x, y)$ if this exists.

For convenience in notation we shall often make use of the function of a rectangle associated with a given function of a point. Thus, if $\rho = (a, b; x, y)$ we shall write $F(\rho) = F(a, b; x, y) = F(x, y) - F(x, b) - F(a, y) + F(a, b)$. We shall say that $F(x, y)$ is in class V if $\sum | F(r_i) |$ is bounded for every non-overlapping set of rectangles $\{r_i\}$ whose sides are parallel to the coördinate axes.[2] If in addition $F(x, \bar{y})$ is of bounded variation in x for at least one value $\bar{y}$ of y, and $F(\bar{x}, y)$ is of bounded variation in y for at least one value $\bar{x}$ of x, we have the definition of the class H mentioned above. Throughout this paper we shall use the term *monotone* to imply a double restriction on $F(x, y)$: we require $F(x', y') \geqq F(x'', y'')$ whenever $x' \geqq x''$, $y' \geqq y''$, and $F(r) \geqq 0$ for every rectangle r.

Now, since functions in class V_1 vanish on the lines $x = a$, $y = b$, while functions in class V need not do so, it is clear that the equation $V = V_1$ cannot be true without modification. We therefore define the classes V_1', V_2' precisely as V_1, V_2, except that (A) is replaced by the more general condition:

$$\text{(A}')\qquad F(a, b; x, y) = \lim_{n\to\infty} \int_{(a,b;x,y)} s_n(x, y)\,dx\,dy.$$

For conciseness we shall denote the class of functions that vanish along the lines $x = a$, $y = b$ by Z. We prove in Theorem 1 below that $H \leqq V_1'$, an obvious corollary of which is $H\cdot Z \leqq V_1$; and since we know from J that $V_1 \leqq H\cdot Z$, it follows that $V_1 = H\cdot Z = V\cdot Z$. Again, it is pointed out by Adams and Clarkson[3] that if $F(x, y)$ is in class V we may write $F(x, y) = \bar{F}(x, y) + G(x) + K(y)$, where $\bar{F}(x, y)$ is in class H. Now $F(a, b; x, y) = \bar{F}(a, b; x, y)$ since for any rectangle r, $G(r) = K(r) = 0$; and since $H \leqq V_1'$, it can be seen by inspection of (A$'$) that $V \leqq V_1'$. But we can prove as in J, Theorem II, that $V_1' \leqq V$. Hence we obtain the result: $V_1' = V > H$.

Lastly we consider functions which are continuous (class C) but not of bounded variation, and prove that $C \leqq V_2'$. It will also be shown that a necessary and sufficient condition for $F(x, y)$ to be in class V_2' is that $F(x, y)$ be in Baire class 0 or 1.

2. In Theorem 1 we shall actually prove more than that $H \leqq V_1'$. We shall show that $s_n(x, y)$ can be defined so that not only does $\int_{(a,b;x,y)} s_n(x, y)\,dx\,dy$

[2] This is equivalent to the definition V given by J. A. Clarkson and C. R. Adams, *On definitions of bounded variation for functions of two variables*, Transactions of the American Mathematical Society, vol. 35(1933), pp. 824–854; p. 825.

[3] C. R. Adams and J. A. Clarkson, *Properties of functions $f(x, y)$ of bounded variation*, Transactions of the American Mathematical Society, vol. 36(1934), pp. 711–730; p. 721.

tend to $F(a, b; x, y)$, but also the integral $\int_{(a,b;x,y)} | s_n(x, y) | \, dx \, dy$ tends to $\int_{(a,b;x,y)} | d_{x,y}F(x,y) |$.[4]

THEOREM 1. *If $F(x, y)$ is in class H, then $s_n(x, y)$ can be defined so that*
(a) *$s_n(x, y)$ is summable,*

(b) $\int_R | s_n(x, y) | \, dx \, dy = \int_R | d_{x,y}F(x, y) |$,

(c) $\lim_{n \to \infty} s_n(x, y) = 0$,

(d) $\lim_{n \to \infty} \int_{(a,b;x,y)} s_n(x, y) \, dx \, dy = F(a, b; x, y)$,

(e) $\lim_{n \to \infty} \int_{(a,b;x,y)} | s_n(x, y) | \, dx \, dy = \int_{(a,b;x,y)} | d_{x,y}F(x, y) |$.

If in addition $F(x, y)$ is continuous, then (d) *and* (e) *can be made to hold uniformly.*

Proof. Since $F(x, y)$ is in class H, there exist functions $P(x, y)$, $N(x, y)$ such that

$$P(x, y) + N(x, y) = \int_{(a,b;x,y)} | d_{x,y}F(x, y) |, \tag{1}$$

$$P(x, y) - N(x, y) = F(a, b; x, y). \tag{2}$$

The functions $P(x, y)$, $N(x, y)$ are monotone, and vanish on the left side and lower side of R.

Let $\xi_1, \xi_2, \cdots$ denote the values of x for which $\alpha(x) = V(x, d) = \int_{(a,b;x,d)} | d_{x,y}F(x, y) |$ is discontinuous, and $\eta_1, \eta_2, \cdots$ the values of y for which $\beta(y) = V(c, y) = \int_{(a,b;c,y)} | d_{x,y}F(x, y) |$ is discontinuous. For each positive integer n select the numbers

$$x_0^n, x_1^n, \cdots, x_h^n; \qquad y_0^n, y_1^n, \cdots, y_k^n,$$

where $h = h(n)$, $k = k(n)$, so that

$$a = x_0^n < x_1^n < \cdots < x_h^n = c,$$

$$b = y_0^n < y_1^n < \cdots < y_k^n = d,$$

$$x_{p+1}^n - x_p^n \leqq n^{-1} \qquad (p = 0, 1, \cdots, h - 1),$$

$$y_{q+1}^n - y_q^n \leqq n^{-1} \qquad (q = 0, 1, \cdots, k - 1);$$

[4] This extension and the proof here given are due to the referee.

also let the numbers $\xi_1, \cdots, \xi_n$ be included among $x_0^n, \cdots, x_h^n$, and $\eta_1, \cdots, \eta_n$ among $y_0^n, \cdots, y_k^n$. Throughout the remainder of this proof we shall understand $p = 0, 1, \cdots, h-1$; $q = 0, 1, \cdots, k-1$.

Denote the open rectangle $(x_p^n, y_q^n; x_{p+1}^n, y_{q+1}^n)$ by r_{pq}^n. In each r_{pq}^n choose arbitrarily the measurable sets $e_{pq}^n, \epsilon_{pq}^n$, so that

$$me_{pq}^n > 0, \qquad m\epsilon_{pq}^n > 0, \qquad e_{pq}^n \cdot \epsilon_{pq}^n = 0,$$
$$\sum_{p,q} me_{pq}^n \leqq 2^{-n}, \qquad \sum_{p,q} m\epsilon_{pq}^n \leqq 2^{-n}.$$

Set

$$s_n'(x, y) = \begin{cases} P(r_{pq}^n)/me_{pq}^n & \text{on } e_{pq}^n, \\ 0 & \text{on } R - \sum_{p,q} e_{pq}^n, \end{cases}$$

$$s_n''(x, y) = \begin{cases} N(r_{pq}^n)/m\epsilon_{pq}^n & \text{on } \epsilon_{pq}^n, \\ 0 & \text{on } R - \sum_{p,q} \epsilon_{pq}^n. \end{cases}$$

Then, except possibly on a set B of measure zero, $s'(x, y)$ and $s''(x, y)$ both tend to zero. Next set

$$s_n^*(x, y) = \begin{cases} 0 & \text{on } B, \\ s_n'(x, y) & \text{on } R - B, \end{cases}$$

$$s_n^{**}(x, y) = \begin{cases} 0 & \text{on } B, \\ s_n''(x, y) & \text{on } R - B. \end{cases}$$

We finally set

$$s_n(x, y) = s_n^*(x, y) - s_n^{**}(x, y),$$

and obtain the sequence required in this theorem. Evidently (a), (b), (c) hold, and we have only to consider (d) and (e).

Let (x, y) be any point of R. We shall show that

$$\lim_{n\to\infty} \int_{(a,b;\,x,y)} s_n^*(x, y)\, dx\, dy = P(x, y). \tag{3}$$

There are four cases to consider, according as x does or does not belong to $\{\xi_i\}$ $(i = 1, 2, \cdots)$ and y does or does not belong to $\{\eta_j\}$ $(j = 0, 1, \cdots)$. Let $X_n' = X_n'(x, y)$, $X_n'' = X_n''(x, y)$ denote numbers $x_\nu^n, x_{\nu+1}^n$ of the set $x_0^n, \cdots, x_h^n$ such that $x_\nu^n \leqq x \leqq x_{\nu+1}^n$. Let Y_n', Y_n'' be similarly defined with respect to the set $y_0^n, \cdots, y_k^n$.

Case I. Suppose $x \in \{\xi_i\}$, $y \in \{\eta_j\}$. Then

$$\int_{(a,b;\,x,y)} s_n^*(x, y)\, dx\, dy = P(x, y)$$

for all sufficiently large values of n. This gives (3) in this case.

Case II. Suppose $x \,\epsilon'\, \{\xi_i\}$, $y \,\epsilon\, \{\eta_j\}$. Then, for all sufficiently large values of n,

$$\begin{aligned}\int_{(a,b;\,x,y)} s_n^*(x, y)\,dx\,dy - P(x, y) &\leqq \int_{(a,b;\,X_n'',y)} s_n^*(x, y)\,dx\,dy - P(x, y)\\ &= P(X_n'', y) - P(x, y)\\ &\leqq P(X_n'', d) - P(x, d)\\ &\leqq \alpha(X_n'') - \alpha(x).\end{aligned}$$

Thus,

$$\overline{\lim_{n\to\infty}} \int_{(a,b;\,x,y)} s_n^*(x, y)\,dx\,dy \leqq P(x, y), \tag{4}$$

since $X_n'' \to x$ as $n \to \infty$, and α is continuous at x. Similarly, using X_n' in place of X_n'', we see that

$$\underline{\lim_{n\to\infty}} \int_{(a,b;\,x,y)} s_n^*(x, y)\,dx\,dy \geqq P(x, y). \tag{5}$$

Equation (3) now follows from (4) and (5).

Case III. If $x \,\epsilon\, \{\xi_i\}$, $y \,\epsilon'\, \{\eta_j\}$, we obtain (3) as in Case II.

Case IV. Suppose $x \,\epsilon'\, \{\xi_i\}$, $y \,\epsilon'\, \{\eta_j\}$. Then, for n sufficiently large,

$$\begin{aligned}\int_{(a,b;\,x,y)} s_n^*(x, y)\,dx\,dy - P(x, y) &\leqq P(X_n'', Y_n'') - P(x, y)\\ &\leqq \alpha(X_n'') - \alpha(x) + \beta(Y_n'') - \beta(y).\end{aligned}$$

Thus (4) holds. Using X_n', Y_n' in place of X_n'', Y_n'', we also see that (5) holds. This gives (3) in the fourth and last case.

Similarly, we can prove that

$$\lim_{n\to\infty} \int_{(a,b;\,x,y)} s_n^{**}(x, y)\,dx\,dy = N(x, y).$$

The required equations (d) and (e) now follow from (1) and (2).

If, finally, $F(x, y)$ is continuous, then $\{\xi_i\}$, $\{\eta_j\}$ are void sets.[5] The uniform continuity of $\alpha(x)$ and $\beta(y)$ gives the uniform approach.

THEOREM 2. *Every function $F(x, y)$ in class V is in class V_1'.*

This follows from Theorem 1 as indicated at the end of the introduction.

THEOREM 3. *Given any function $F(x, y)$ in class V, and any summable function $f(x, y)$, we can find a sequence $s_n(x, y)$ such that $s_n(x, y) \to f(x, y)$, $\int_{(a,b;\,x,y)} s_n(x, y)\,dx\,dy \to F(a, b; x, y)$, and $\int_e s_n(x, y)\,dx\,dy$ is bounded in n and e.*

[5] See J. Pierpont, *Theory of Functions of Real Variables*, New York, 1912, vol. II, p. 531, where the proof is given for a function $f(x)$ of a single real variable. It is easily modified to give the result required here.

Proof. The function $G(x, y) = F(x, y) - \int_{(a,b;x,y)} f(x, y)\, dx\, dy$ is in class V. We can therefore determine a sequence $t_n(x, y)$ such that $t_n(x, y)$ tends to zero everywhere, $\int_{(a,b;x,y)} t_n(x, y)\, dx\, dy$ tends to $G(a, b; x, y)$, and $\int_e t_n(x, y)\, dx\, dy$ is bounded in n and e. Then the sequence $s_n(x, y) = t_n(x, y) + f(x, y)$ satisfies the requirements of the present theorem.

These results are of interest in connection with a theorem of Jeffery,[6] that a necessary and sufficient condition for a function $F(x)$ to be of bounded variation in the interval (a, b) is that there exist a sequence of summable functions $s_n(x)$ such that $s_n(x)$ tends to $F'(x)$ where this is defined, $\int_a^x s_n(x)\, dx$ tends to $F(x) - F(a)$, and $\int_e s_n(x, y)\, dx$ is bounded in n and e. From the results of the present paper we obtain the following extension of this theorem to functions of two variables: A necessary and sufficient condition for a function $F(x, y)$ to be of bounded variation, in the sense that $\sum | F(r_i) |$ is bounded, is that there exist a sequence of summable functions $s_n(x, y)$ such that $s_n(x, y)$ tends to $F'(x, y)$ where this is defined, $\int_{(a,b;x,y)} s_n(x, y)\, dx\, dy$ tends to $F(a, b; x, y)$, and $\int_e s_n(x, y)\, dx\, dy$ is bounded in n and e. In this statement we may replace "$F'(x, y)$ where this is defined" by "an arbitrary summable function $f(x, y)$"; so also for a function of a single variable.

3. We now consider functions which are continuous but not in class V. It is convenient to prove first the following lemma.

LEMMA. *Let $F(x, y)$ be a continuous function. Then for each positive integer n we can find a function $S_n(x, y)$ such that*

(a) *$S_n(x, y)$ is summable,*

(b) $mE\{(x, y)\colon S_n(x, y) \neq 0\} \leqq 1/2^n$,

(c) $\left| F(a, b; x, y) - \int_{(a,b;x,y)} S_n(x, y)\, dx\, dy \right| \leqq 9/n$.

Proof. We use a construction similar to that of Theorem 1. Divide $[a, c]$ into $h = h(n)$ equal parts by the points $a = x_0^n < x_1^n < \cdots < x_h^n = c$, and $[b, d]$ into $k = k(n)$ equal parts by the points $b = y_0^n < y_1^n < \cdots < y_k^n = d$, in such a way that the oscillation of $F(x, y)$ in each rectangle $r_{pq}^n\colon(x_p^n, y_q^n; x_{p+1}^n, y_{q+1}^n) \leqq 1/n$. As in Theorem 1, $p = 0, 1, \cdots, h - 1$; $q = 0, 1, \cdots,$

[6] R. L. Jeffery, *Functions defined by sequences of integrals and the inversion of approximate derived numbers*, Transactions of the American Mathematical Society, vol. 41(1937), pp. 171–192; p. 175.

$k - 1$. In each rectangle r^n_{pq} choose a set e^n_{pq} so that $me^n_{pq} > 0$, $\sum_{p,q} me^n_{pq} \leqq 2^{-n}$, the sets e^n_{pq} being obtained by translating e^n_{00} through multiples of Δx, Δy, the common lengths of the x and y subintervals.

We then define $S_n(x, y) = F(r^n_{pq})/me^n_{pq}$ on e^n_{pq}, $S_n(x, y) = 0$ on $R - \sum_{p,q} e^n_{pq}$. Then (a) and (b) obviously hold. To prove (c), we define X'_n, Y'_n, X''_n, Y''_n as in Theorem 1, and write

$$|F(a, b; x, y) - F(a, b; X'_n, Y'_n)| \leqq \frac{3}{n},$$

$$F(a, b; X'_n, Y'_n) = \int_{(a,b;X'_n,Y'_n)} S_n(x, y)\,dx\,dy,$$

$$\left|\int_{(X'_n,Y'_n;x,y)} S_n(x, y)\,dx\,dy\right| \leqq |F(X'_n, Y'_n; X''_n, Y''_n)| \leqq \frac{2}{n},$$

$$\left|\int_{(X'_n,b;x,Y'_n)} S_n(x, y)\,dx\,dy\right| \leqq |F(X'_n, b; X''_n, Y'_n)| \leqq \frac{2}{n},$$

$$\left|\int_{(a,Y'_n;X'_n,y)} S_n(x, y)\,dx\,dy\right| \leqq |F(a, Y'_n; X'_n, Y''_n)| \leqq \frac{2}{n}.$$

The lemma follows on combining these inequalities.

THEOREM 4. *Let $F(x, y)$ be continuous. Then we can find a sequence of functions $s_n(x, y)$ such that $s_n(x, y)$ tends to zero everywhere, and the integral $\int_{(a,b;x,y)} s_n(x, y)\,dx\,dy$ tends uniformly to $F(a, b; x, y)$.*

Proof. Except possibly on a set B of measure zero, the functions $S_n(x, y)$ of the lemma tend to zero. On setting $s_n(x, y) = 0$ on B, $s_n(x, y) = S_n(x, y)$ on $R - B$, we obtain the sequence required in this theorem.

THEOREM 5. *Let $F(x, y)$ be continuous, and let $f(x, y)$ be any measurable function. Then we can find a sequence of functions $s_n(x, y)$ such that $s_n(x, y)$ tends to $f(x, y)$, and $\int_{(a,b;x,y)} s_n(x, y)\,dx\,dy$ tends uniformly to $F(a, b; x, y)$.*

Proof. Set

$$f_n(x, y) = \begin{cases} f(x, y), & |f(x, y)| \leqq n, \\ n \operatorname{sgn} f(x, y), & |f(x, y)| > n. \end{cases}$$

Then for each value of n the function

$$G_n(x, y) = F(x, y) - \int_{(a,b;x,y)} f_n(x, y)\,dx\,dy$$

is continuous. Determine as in the lemma a sequence $T_n(x, y)$ such that

$$\left| G_n(a, b; x, y) - \int_{(a,b;x,y)} T_n(x, y)\, dx\, dy \right| \leqq \frac{9}{n}.$$

Then $S_n(x, y) = T_n(x, y) + f_n(x, y)$ satisfies the inequality

$$\left| F(a, b; x, y) - \int_{(a,b;x,y)} S_n(x, y)\, dx\, dy \right| \leqq \frac{9}{n},$$

and $S_n(x, y)$ tends to $f(x, y)$ except possibly on a set B of measure zero. Finally we set $s_n(x, y) = f(x, y)$ on B, $s_n(x, y) = S_n(x, y)$ on $R - B$, and obtain the required sequence.

We have thus proved that every continuous function is in class V_2', and that we may take any measurable function whatever as the function $f(x, y)$ which occurs in the definition of this class. In conclusion we state the following theorem.

THEOREM 6. *A necessary and sufficient condition that $F(x, y)$ be in class V_2' is that $F(x, y)$ be in Baire class* 0 *or* 1 (*i.e., that $F(x, y)$ be continuous or the limit of a sequence of continuous functions*).

The necessity of the condition is obvious. To prove the sufficiency, in the case where $F(x, y) = \lim_{n\to\infty} F_n(x, y)$, say, we need only determine $S_n(x, y)$ as in the lemma, so that

$$\left| F_n(a, b; x, y) - \int_{(a,b;x,y)} S_n(x, y)\, dx\, dy \right| \leqq \frac{9}{n},$$

then take $s_n(x, y)$ as in Theorem 4.

ACADIA UNIVERSITY.

INTERVAL-FUNCTIONS

By André Gleyzal

Introduction. Let T be a given set of open intervals I of the continuum, and let $\varphi(I)$ be a real function on T; we then say $\varphi(I)$ is an *interval-function* on the *range* T. We shall consider such interval-functions and point-functions readily associated with them. Certain of these point-functions are shown to be characterizable as upper or lower semi-continuous or as the monotone limit of a sequence of such functions. One of the associated point-functions is what we term the *kernel* of the interval-function. In general, this kernel is many-valued; of particular interest is the case where it is one-valued. We then call the interval-function *convergent*. It is shown that a necessary and sufficient condition for a point-function to be the kernel of a convergent interval-function is that it be the limit of a sequence of continuous functions. We prove furthermore, without reference to Baire's theorem, that the kernel of a convergent interval-function has a point of continuity on every perfect set with respect to the set;[1] and conversely, if a function $f(x)$ has a point of continuity on every perfect set with respect to the set, it is the kernel of a convergent interval-function. These two necessary and sufficient conditions yield a proof of the Baire theorem along lines different from those in the literature.

These results may be applied, for example, when the interval-function is chosen as follows. Let $f(x)$ be a given real function defined, say, on the entire continuum. We choose as the number $\varphi(I)$ associated with an interval $I = (a, b)$ the difference quotient $(f(b) - f(a))/(b - a)$. Allowing I to vary over the set of all intervals, we secure an interval-function $\varphi(I)$. When $f'(x)$ exists, $\varphi(I)$ is what we have termed convergent and $f'(x)$ is the kernel of $\varphi(I)$. Again, $\varphi(I)$ may be chosen as the upper boundary of a function $f(x)$ in the interval I, or as the saltus of $f(x)$ in I divided by the length of I. Other interval-functions may be obtained in a similar manner from a real function $f(x)$ by neglecting the values of $f(x)$ on sets which are negligible from the point of view of cardinal number, Lebesgue measure, etc. A number of interval-functions arise, also, in the study of a point-set S. For example, we may associate with an interval I the exterior Jordan or Lebesgue measure of S in I, or the relative exterior

Received June 21, 1940. The author wishes to express his thanks to Professor H. Blumberg for criticisms and suggestions during the preparation of this paper.

[1] We prove that such a point of continuity exists in the interior of the perfect set, i.e., the point is neither the left nor the right end-point of the set. In another paper *On the equation* $dy/dx = f(x, y)$, Bulletin of the American Mathematical Society, vol. 47, pp. 254–256, the author requires this form of the theorem.

measure of S in I. In the latter case, if the interval-function is convergent, the value of its kernel at a point ξ is equal to the metric density of S at ξ.

1. **Characterization of $L(x)$ and $U(x)$.** An interval-function $\varphi(I)$ defined on a set of open intervals $T = \{I\}$ may be geometrically represented by means of a set of line segments $T' = \{I'\}$ of the xy-plane, the segment I', equal in length to I, being placed parallel to the interval I at distance $|\varphi(I)|$ directly above or below it according as $\varphi(I)$ is positive or negative.

Suppose there is given an interval-function $\varphi(I)$ defined on a range $T = \{I\}$. To every point x there corresponds a set of intervals $\{I_x\}$ containing x, hence a set of numbers $\{\varphi(I_x)\}$. Let us denote the latter set by $\Phi(x)$. In general, $\Phi(x)$ is a many-valued function. The set of numbers $\Phi(x)$, for a particular x, has a lower boundary $L(x)$ and an upper boundary $U(x)$ which we shall call, respectively, the lower and upper point-function associated with $\varphi(I)$.

We prove that $L(x)$ is upper semi-continuous in S—the set of all points belonging to intervals I of T. For let ξ be a point of S and ϵ any positive number. Our definition of $L(\xi)$ implies the existence of an interval I_ξ of T containing ξ such that $\varphi(I_\xi) \leqq L(\xi) + \epsilon$. If x is any point in I_ξ, $L(x) \leqq \varphi(I_\xi)$ and hence $L(x) \leqq L(\xi) + \epsilon$. Therefore $L(x)$ is upper semi-continuous.

Conversely, given an upper semi-continuous function $f(x)$ defined on an open interval (a, b), there exists an interval-function $\varphi(I)$ such that $f(x)$ is the lower point-function associated with $\varphi(I)$. To prove this, we set $\varphi(I)$ equal to the upper boundary of $f(x)$ for x in I. If ξ is any point of (a, b), we may write $f(\xi) \leqq \varphi(I)$ for all intervals I containing ξ. Therefore $f(\xi) \leqq L(\xi)$, where $L(\xi)$ is the lower point-function of $\varphi(I)$. Also, given any positive number ϵ, there exists an interval I containing ξ such that $\varphi(I) < f(\xi) + \epsilon$, since $f(x)$ is upper semi-continuous. But $L(\xi) < \varphi(I)$ and consequently $L(\xi) < f(\xi) + \epsilon$. Hence $L(\xi) \leqq f(\xi)$ and we conclude $L(\xi) = f(\xi)$.

2. **Characterization of $u(x)$ and $l(x)$.** Suppose $\varphi(I)$ is a given interval-function defined on a set of intervals T. By the kernel K of T we understand the set of points x covered by an interval of T of infinitesimal length. We associate with $\varphi(I)$ the function $\varphi(x)$ (in general many-valued) obtained from $\varphi(I)$ in the following manner. Let ξ be a point of K. $\varphi(\xi)$ is defined to be the set of limiting values (we permit inclusion of $\pm\infty$) of $\varphi(I)$, where I is an interval containing ξ of infinitesimal length. More explicitly, a real number η is an element of the set of numbers $\varphi(\xi)$ if there exists a sequence $\{I_n\}$ of intervals containing ξ and of length approaching zero as n goes to infinity, such that $\eta = \lim_{n\to\infty} \varphi(I_n)$. We call $\varphi(x)$ the *kernel* of $\varphi(I)$. If ξ is any point of K, the set of values $\varphi(\xi)$ has an upper boundary $u(\xi)$ and a lower boundary $l(\xi)$. If we allow ξ to vary, there result two functions $u(x)$ and $l(x)$ defined on K; we call these respectively the *lower kernel* and the *upper kernel* of $\varphi(I)$.

THEOREM. *The lower kernel of an interval-function is the monotone (upward) limit of a sequence of upper semi-continuous functions, and conversely, every func-*

tion which is the limit of a sequence of upper semi-continuous functions is th lower kernel of an interval-function. A corresponding statement holds for the uppe kernel of an interval-function.

Proof. Suppose a range T and an interval-function $\varphi(I)$ defined on T ar given. Let T_n be the set of intervals of T of length less than n^{-1}, where n is a positive integer; ξ a point of the kernel K of T; and $L_n(\xi)$ the lower boundary of values $\varphi(I)$, for intervals I of T_n which contain ξ. As ξ varies we obtain a function $L_n(x)$. This function, as we have shown, is upper semi-continuous. In addition, it is clear that $L_n(\xi) \leqq L_{n+1}(\xi)$ for all n. Suppose $l(\xi)$ (the value at ξ of the kernel of $\varphi(I)$) is finite. Obviously $L_n(\xi) \leqq l(\xi)$ for every integer n. Also, given a positive number ϵ, the definitional properties of $l(\xi)$ imply the existence of an integer m such that $l(\xi) < \varphi(I) + \epsilon$ for all I of T_m containing ξ. Therefore, $l(\xi) \leqq L_n(\xi) + \epsilon$ for all n greater than or equal to m. Consequently, $l(\xi) = \lim_{n\to\infty} L_n(\xi)$. Suppose, now, ξ is a point of K where $l(\xi)$ has the value $-\infty$. Then $L_n(\xi) = -\infty$ for all n and $\lim_{n\to\infty} L_n(\xi) = l(\xi)$. Finally, if $l(\xi) = +\infty$, then for every positive number N the definitional properties of $l(\xi)$ insure the existence of an integer n such that $\varphi(I) > N$ for all I of T_n containing ξ. It follows that $L_n(\xi) > N$, and we have again $l(\xi) = \lim_{n\to\infty} L_n(\xi)$. We conclude that $l(\xi)$ is the monotone limit of a sequence of upper semi-continuous functions.

We now prove that, conversely, a function $f(x)$ which is the limit of a sequence of upper semi-continuous functions $\{f_n(x)\}$ is equal to the lower kernel $l(x)$ of some interval-function $\varphi(I)$. We suppose $f(x)$ defined on an open interval (a, b). Let us assume, first, that $f(x)$ is bounded above and below. If k is a given integer, $f_k(x)$ has a lower boundary l_{k1} in (a, b) and there exists a point ξ_1 of (a, b) such that $f_k(\xi_1)$ is less than $l_{k1} + k^{-1}$. Since $f_k(x)$ is upper semi-continuous, there exists an interval I_{k1} containing ξ_1 of length less than k^{-1} such that $f_k(x)$, for values x in I_{k1}, is less than $l_{k1} + k^{-1}$. As a first step in the definition of $\varphi(I)$, we set $\varphi(I_{k1}) = l_{k1} + k^{-1}$. Let l_{k2} be the lower boundary of $f_k(x)$ for values x in $(a, b) - I_{k1}$. There exists, as before, an interval I_{k2} of length less than k^{-1} such that $f_k(x)$, x in I_{k2}, is less than $l_{k2} + k^{-1}$. We set $\varphi(I_{k2}) = l_{k2} + k^{-1}$. Suppose $I_{k\lambda}$ is defined for all ordinals $\lambda < \mu$. In a similar manner we associate with the set of points $(a, b) - (I_{k1} + I_{k2} + \cdots + I_{k\omega} + \cdots + I_{k\lambda} + \cdots)$, $\lambda < \mu$, an interval $I_{k\mu}$ and a value $\varphi(I_{k\mu})$.

We show that the lower kernel $l(x)$ of $\varphi(I)$ as thus defined is equal to $f(x)$. Let k be a given positive integer, and ξ a point of (a, b). There will be a first interval $I_{k\mu}$ of the ordered set $I_{k1}, I_{k2}, \cdots; I_{k\omega}, \cdots, I_{k\lambda}, \cdots$ containing ξ. $I_{k\mu}$ is of length less than k^{-1} and $\varphi(I_{k\mu})$ differs from $f_k(\xi)$ by less than k^{-1}. Consequently $f(\xi)$ is a limiting value for $\varphi(I)$, I infinitesimal, on ξ. Therefore, $l(\xi) \leqq f(\xi)$. We prove, now, that $l(\xi) \geqq f(\xi)$. Let ϵ be a small positive number. We may choose an integer m sufficiently large to insure that the saltus of values $f_k(\xi)$, for $k > m$, is less than ϵ. With every integer k, there is associated, as we have seen, an interval $I_{k\mu}$. We note that no interval preceding $I_{k\mu}$ in the order $I_{k1}, I_{k2}, \cdots; I_{k\omega}, \cdots, I_{k\lambda}, \cdots$ contains ξ. Also, the intervals coming after

$I_{k\mu}$ which contain ξ have length greater than a fixed number η_k since such an interval also contains a point not in $I_{k\mu}$. We set η equal to the least of the numbers $\eta_1, \cdots, \eta_{m-1}$. Since we are seeking information concerning infinitesimal intervals, we may disregard intervals $I_{k\lambda}$ of length greater than or equal to η. We thus consider only intervals $I_{k\lambda}$ ($k \geqq m$) containing ξ. But values $\varphi(I_{k\lambda})$ for such intervals are greater than $f_k(\xi)$. Since $|f_k(\xi) - f(\xi)| < \epsilon$ if $k > m$, it follows that $\varphi(I_{k\mu}) > f(\xi) - \epsilon$ for intervals $I_{k\lambda}$ where $k > m$. Consequently, if ϵ is any number, $\varphi(I_{k\mu}) > f(\xi) - \epsilon$ for sufficiently small intervals $I_{k\lambda}$ containing ξ. Therefore $L(\xi) \geqq f(\xi)$, and we conclude that $L(\xi) = f(\xi)$.

Suppose now that $f(x)$ is unbounded. The proof of the theorem in this case may be made to depend on the one we have just considered by means of the transformation $y^* = y/(1 + |y|)$. For $f^*(x) = f(x)/(1 + |f(x)|)$, the transform of $f(x)$, is never greater than unity in absolute value and is therefore the lower kernel of an interval-function $\varphi^*(I)$. Consideration of the construction of $\varphi^*(I)$ enables us to verify the possibility of choosing the values $\varphi^*(I)$ to be not greater than unity in absolute value. Thus $\varphi^*(I)$ may be regarded as the transform of an interval-function $\varphi(I) = \varphi^*(I)/(1 - |\varphi^*(I)|)$, and we conclude $f(x)$ is the lower kernel of $\varphi(I)$.

The proof of the corresponding statement for the upper kernel is obtained simply by making substitutions such as *lower* for *upper*, etc.

Given $f(x)$, the limit of a sequence of upper semi-continuous functions, we may construct the interval-function $\varphi(I)$ as above. $f(x)$ is therefore the monotone (upward) limit of a sequence of upper semi-continuous functions $\{L_n(x)\}$ since it is the lower kernel of $\varphi(I)$. Thus a function which is the limit of a sequence of upper semi-continuous functions is the monotone (upward) limit of a sequence of upper semi-continuous functions.

If a function $f(x)$ is the limit of a sequence of upper semi-continuous functions and also the limit of a sequence of lower semi-continuous functions, it is the limit of a sequence of continuous functions. For $f(x)$ is the monotone (upward) limit of a sequence of upper semi-continuous functions $\{L_n(x)\}$ and the monotone (downward) limit of a sequence of lower semi-continuous functions $\{U_n(x)\}$. By the *Einschiebungs theorem* a continuous function $f_n(x)$ may be inserted between the two functions $L_n(x)$ and $U_n(x)$. It follows that $f(x) = \lim_{n\to\infty} f_n(x)$.

3. **Proof of the Baire theorem.** If an interval-function $\varphi(I)$ is such that its many-valued kernel $\varphi(x)$ is one-valued everywhere, we shall call $\varphi(I)$ a *convergent* interval-function. In this case the lower and upper kernels $l(x)$ and $u(x)$ of $\varphi(I)$ are equal to $\varphi(x)$. The *Einschiebungs theorem* tells us, then, that $\varphi(x)$ is the limit of a sequence of continuous functions. We shall prove the latter statement and its converse without reference to this theorem.

THEOREM. *The kernel $\varphi(x)$ of a convergent interval-function is the limit of a sequence of continuous functions, and conversely, if a function $f(x)$ is the limit of a sequence of continuous functions, a convergent interval-function exists such that its kernel is identical with $f(x)$.*

Proof. Let $\varphi(I)$ be a convergent interval-function, the argument I ranging over a set of open intervals T. In terms of $\varphi(I)$ we define a sequence of continuous functions $\{f_n(x)\}$ as follows. For every integer n there exists a subset of intervals T_n of T, each of length less than n^{-1}, such that every point of K, the kernel of T, is covered by some interval of T_n and not more than two intervals of T_n cover any single point of K. We may assume, too, that each interval of T_n contains a point of K. If there is but one interval I of T_n containing a point x of K, let $f_n(x) = \varphi(I)$. If there are two intervals I_1 and I_2 of T_n containing a point x of K, there will be an open interval I_1I_2 containing x, and we define $f_n(x)$ linearly on I_1I_2, setting $f_n(x)$ equal to $\varphi(I_1)$ at the left end-point of I_1I_2 and to $\varphi(I_2)$ at the right end-point. We have thus defined a continuous function $f_n(x)$. If x is a given point of K and n a sufficiently large integer, $\varphi(x)$, the kernel of $\varphi(I)$, differs from $\varphi(I)$ by as little as we please for intervals I containing x of length less than n^{-1}. Consequently, $\varphi(x)$ differs from $f_n(x)$ by as little as desired provided n is chosen large enough. Therefore $\lim_{n\to\infty} f_n(x) = \varphi(x)$.

Conversely, suppose $f(x)$—defined on an open interval (a, b)—is the limit of a sequence of continuous functions $\{f_n(x)\}$. We construct a convergent interval-function $\varphi(I)$ such that its kernel equals $f(x)$. If n is a positive integer, there exists a set of open intervals $T_n = \sum_{k=1}^{\infty} I_{nk}$ which cover (a, b), each interval I_{nk} being of length less than n^{-1}, such that the saltus of $f_n(x)$, for x in any interval I_{nk}, is less than n^{-1}. We may assume that not more than two intervals of T_n cover any one point of S. We choose as the range T of $\varphi(I)$ the set of all intervals I_{nk} $(1 < n < k, 1 < k < \infty)$ and set $\varphi(I_{nk}) = f_n(x)$, where x is the abscissa of the mid-point of I_{nk}. Thus $\varphi(I)$ is defined on T.

Consider any point ξ of (a, b). Given an integer k, if l is sufficiently large, the saltus of values $f_n(\xi), f(\xi)$ for $n > l$ is less than k^{-1}. Let $\{I_n\}$ be a sequence of intervals of T, each containing ξ and of vanishing length as $n \to \infty$. Each interval I_n belongs to some set T_m; but not more than two belong to one T_m. Therefore, corresponding to the integer k there exists also an integer l' such that I_n, for all $n > l'$, belongs to sets T_m where $m > k$. The saltus of $\varphi(I), f(\xi)$ for intervals I of the set $\sum_{n=l+l'}^{\infty} I_n$ is less than $4k^{-1}$. Thus $\varphi(I)$ is a convergent interval-function, and its kernel is equal to $f(x)$.

As a second step in the proof of the Baire theorem we establish the following property of $\varphi(x)$.

THEOREM. *If $\varphi(x)$ is the kernel of a convergent interval-function and S is a perfect set, there exists a point of S interior*[2] *to S which is a point of continuity of $\varphi(x)$ with respect to S; conversely, if a function $f(x)$ defined on an open interval has a point of continuity on every perfect set S with respect to S, a convergent interval-function exists such that its kernel equals $f(x)$.*

[2] See footnote 1.

Proof. Let $\varphi(I)$ be a convergent interval-function defined on a range T. Suppose that for a given perfect set S the saltus of its kernel $\varphi(x)$ for points x in SI is greater than a fixed positive ϵ for every closed interval I containing points of S.[3] We show that this assumption leads to a contradiction. Let I_1 be an interval of T such that $SI_1 \neq 0$. There exist, then, two points ξ_1 and ξ_2 of SI_1 such that $|\varphi(\xi_2) - \varphi(\xi_1)| > \epsilon$. We may therefore find intervals J and J' interior to I_1 such that $SJ \neq 0$, $SJ' \neq 0$, and $|\varphi(J) - \varphi(J')| > \epsilon$. One of the two values $\varphi(J)$, $\varphi(J')$ differs from $\varphi(I_1)$ by more than $\frac{1}{2}\epsilon$. Let J_1 be the interval J or J' for which this is true. Then $|\varphi(I_1) - \varphi(J_1)| > \frac{1}{2}\epsilon$. We start anew with an interval I_2 of T whose end-points are interior to J_1, of length less than one-half the length of I_1, and such that $SI_2 \neq 0$. We may show, as before, that there exists an interval J_2 of T interior to I_2 such that $SJ_2 \neq 0$ and $|\varphi(I_2) - \varphi(J_2)| > \frac{1}{2}\epsilon$. Similarly, we define the intervals I_n, J_n for all positive integers n. Let ξ be the point common to all the intervals I_n and J_n. $\varphi(I)$ cannot be convergent at ξ since $|\varphi(I_n) - \varphi(J_n)| > \frac{1}{2}\epsilon$ for all n.

Having reached a contradiction, we may conclude that for every perfect set S there exists a closed interval I_1 containing a point of S, such that the saltus of $\varphi(x)$ in I_1 with respect to S does not exceed ϵ, an arbitrarily chosen, small positive number. We assume, as we may, the end-points of I_1 to be interior points of S. SI_1 is a perfect set and again it is possible to find a closed interval I_2 interior to I_1 (end-points also interior) containing a point of S and of length less than half that of I_1, such that the saltus of $\varphi(x)$ in I_2 with respect to S does not exceed $\frac{1}{2}\epsilon$. Similarly, we determine $I_3, I_4, \cdots, I_n, \cdots$ for all integers n. The point x common to all the intervals I_n is a point of S interior to S. Moreover, x is a point of continuity of $\varphi(x)$ with respect to S and the first statement of the theorem is established.

We now prove that if $f(x)$ has a point of continuity on every perfect set S with respect to S, it is the kernel of a convergent interval-function. We assume $f(x)$ to be defined on an open interval (a, b) and construct the desired interval-function $\varphi(I)$. The closed interval Π consisting of (a, b) plus end-points is a perfect set and consequently possesses a point of continuity of f. There exists, therefore, an open interval I of (a, b) such that the saltus of f in I is less than ϵ, an arbitrarily chosen small positive number. Let I_1 be a subinterval of I of length less than ϵ. We choose $\varphi(I_1)$ of such value that for every x of I_1

$$|\varphi(I_1) - f(x)| < \epsilon.$$

For convenience, we shall write an inequality such as the latter in the form $\varphi(I) \doteq f(x)$. Suppose the intervals I_ν are defined for all ordinals ν less than an ordinal λ. Either every point of (a, b) is contained in $\sum = \sum_{\nu<\lambda} I_\nu$, or there is a closed set $\Pi - \Sigma$ remaining which contains points of (a, b). If $(a, b) - \Sigma$ has an isolated point ξ, let I_λ, of length less than ϵ and interior to (a, b),

[3] If $\varphi(x)$ is defined for only one or no point of SI, the statement that $\varphi(x)$ has a point of continuity on S with respect to S is trivially or vacuously true.

contain this point but no other point of $(a, b) - \Sigma$. We set $\varphi(I_\lambda) = f(\xi)$. If $(a, b) - \Sigma$ has no isolated point, it is a perfect set except possibly for omission of end-points. In any case, f has a point of continuity on $\Pi - \Sigma$ with respect to $\Pi - \Sigma$. It is therefore possible to find an interval I_λ of (a, b) of length less than ϵ, such that the saltus of $f(x)$ for x in $I_\lambda[(a, b) - \Sigma]$ is less than ϵ. We choose $\varphi(I_\lambda)$ so that $\varphi(I_\lambda) \overset{\epsilon}{=} f(x)$ for x in $I_\lambda[(a, b) - \Sigma]$.

Only a denumerable number of I's will be needed to cover (a, b) since each I_λ covers a new point of (a, b). The intervals I_λ, which we shall term ϵ-intervals, have the property that $\varphi(I_{\lambda_1}) \overset{\epsilon}{=} f(x)$ if x is in the set $S_{\lambda_1} = I_{\lambda_1} \cdot \{(a, b) - \sum_{\nu_1 < \lambda_1} I_{\nu_1}\}$. In the above equations, for purposes of uniform notation, we have substituted ν_1 for ν and λ_1 for λ as current ordinals.

We now define $\varphi(I)$ on a set of $\epsilon/2$-intervals. Consider the set S_{λ_1}. If it has an isolated point ξ, there exists an interval $I_{\lambda_1 1}$ interior to I_{λ_1} which contains ξ but no other point of S_{λ_1} and is of length less than $\frac{1}{2}\epsilon$. Let $\varphi(I_{\lambda_1 1}) = f(\xi)$. If S_{λ_1} has no isolated point, it is perfect except possibly for omission of end-points. As before, we may find an interval $I_{\lambda_1 1}$ of length less than $\frac{1}{2}\epsilon$ interior to I_{λ_1} containing points of S_{λ_1}, such that the saltus of $f(x)$ for points x in $S_{\lambda_1} I_{\lambda_1 1}$ is less than $\frac{1}{2}\epsilon$. We choose $\varphi(I_{\lambda_1 1})$ so that $\varphi(I_{\lambda_1 1}) \overset{\frac{1}{2}\epsilon}{=} f(x)$, x in $S_{\lambda_1} I_{\lambda_1 1}$. Starting with $S_{\lambda_1} - I_{\lambda_1 1}$ instead of S_{λ_1}, we may find an interval $I_{\lambda_1 2}$ of length less than $\frac{1}{2}\epsilon$ and a value $\varphi(I_{\lambda_1 2})$ such that $\varphi(I_{\lambda_1 2}) \overset{\frac{1}{2}\epsilon}{=} f(x)$ for points x in $I_{\lambda_1 2} \cdot (S_{\lambda_1} - I_{\lambda_1 1})$. Similarly, we determine $I_{\lambda_1 3}, I_{\lambda_1 4}, \cdots ; I_{\lambda_1 \omega}, \cdots, I_{\lambda_1 \lambda}, \cdots$ and the functional values φ until all points of S_{λ_1} are contained in some $I_{\lambda_1 \lambda}$. We shall call these intervals $\epsilon/2$-intervals. They have the property that $\varphi(I_{\lambda_1\lambda_2}) \overset{\frac{1}{2}\epsilon}{=} f(x)$ if x is in the set

$$S_{\lambda_1\lambda_2} = I_{\lambda_1\lambda_2} \cdot \{S_{\lambda_1} - \sum_{\nu_2 < \lambda_2} I_{\lambda_1\nu_2}\}.$$

We now take each $S_{\lambda_1\lambda_2}$ and determine, in a similar manner, $\epsilon/3$-intervals $I_{\lambda_1\lambda_2}{}^{\lambda}$ interior to $I_{\lambda_1\lambda_2}$ and corresponding functional values φ. Continuing, we determine a set of ϵ/n-intervals $I_{\lambda_1\cdots\lambda_n}$ and $\varphi(I_{\lambda_1\cdots\lambda_n})$ for all positive integers n. We may write the relations: $\varphi(I_{\lambda_1\cdots\lambda_n}) \overset{\epsilon/n}{=} f(x)$ if x is in the set

$$S_{\lambda_1\cdots\lambda_n} = I_{\lambda_1\cdots\lambda_n}\{S_{\lambda_1\cdots\lambda_{n-1}} - \sum_{\nu_n < \lambda_n} I_{\lambda_1\cdots\lambda_{n-1}\nu_n}\}.$$

This completes the definition of $\varphi(I)$.

We show that $\varphi(I)$, for I's containing a fixed point ξ of (a, b), has the unique limit $f(\xi)$ as the lengths of the I's approach zero. There exists a first ϵ-interval I_{α_1} of the normal order $I_1, I_2, \cdots ; I_\omega, \cdots, I_\lambda, \cdots$ containing ξ. The intervals $I_{\lambda_1\cdots\lambda_n}$ ($\lambda_1 < \alpha_1$; $n = 1, 2, 3, \cdots$) do not contain ξ and we are not concerned with them. The intervals $I_{\lambda_1\cdots\lambda_n}$ ($\lambda_1 > \alpha_1$; $n = 1, 2, 3, \cdots$) which contain ξ are of length greater than a fixed positive number η_1 ; for these intervals contain ξ and also a point of $(a, b) - I_{\alpha_1}$, since they were defined upon considering points not in I_{α_1}. We need consider, then, only the φ's for the intervals $I_{\alpha_1\lambda_2\cdots\lambda_n}$ ($n = 2, 3, 4, \cdots$). We have $\varphi(I_{\alpha_1}) \overset{\epsilon}{=} f(\xi)$; also,

$$\varphi(I_{\alpha_1\lambda_2\cdots\lambda_n}) \overset{\epsilon+\frac{1}{2}\epsilon}{=} f(\xi) \qquad (n = 2, 3, 4, \cdots),$$

since these functional values approximate some value of $f(x)$, x in S_{α_1}, by less than $\frac{1}{2}\epsilon$, and $f(x)$ has saltus less than ϵ in S_{α_1}. Again, there exists a first $I_{\alpha_1\alpha_2}$ of the normal order $I_{\alpha_1 1}, I_{\alpha_1 2}, \cdots; I_{\alpha_1\omega}, \cdots, I_{\alpha_1\lambda}, \cdots$ containing ξ. We are not concerned with $I_{\alpha_1\lambda_1\cdots\lambda_n}$ ($\lambda_2 < \alpha_2$; $n = 2, 3, 4, \cdots$), since they do not contain ξ. The intervals $I_{\alpha_1\lambda_2\cdots\lambda_n}$ ($\lambda_2 > \alpha_2$; $n = 2, 3, 4, \cdots$) containing ξ also contain some point of $(a, b) - I_{\alpha_1\alpha_2}$ and are therefore all of length greater than some fixed positive number η_2. As before, $\varphi(I_{\alpha_1\alpha_2}) \stackrel{\frac{1}{3}\epsilon}{=} f(\xi)$ and all values $\varphi(I_{\alpha_1\alpha_2\lambda_3\cdots\lambda_n})$ ($n = 3, 4, 5, \cdots$) are within $\frac{1}{2}\epsilon + \frac{1}{3}\epsilon$ of $f(\xi)$. Continuing, we obtain a definite sequence of numbers $\eta_1, \eta_2, \eta_3, \cdots$. Let us consider only intervals containing ξ and of length less than η_r', where η_r' is the least of the numbers $\eta_1, \eta_2, \cdots, \eta_r$. This eliminates the intervals $I_{\lambda_1\cdots\lambda_n}$ ($\lambda_1 \neq \alpha_1$; $n = 1, 2, 3, \cdots$); also, $I_{\alpha_1\lambda_2\cdots\lambda_n}$ ($\lambda_2 \neq \alpha_2$; $n = 2, 3, 4, \cdots$), etc., leaving the intervals $I_{\alpha_1\cdots\alpha_r}$ and $I_{\alpha_1\cdots\alpha_r\lambda_{r+1}\cdots\lambda_n}$ ($n = r + 1, r + 2, r + 3, \cdots$) to consider. The values of φ for these intervals differ from $f(\xi)$ by less than $\epsilon/r + \epsilon/(r + 1)$. Since r is an arbitrarily large integer, we conclude $\varphi(I)$ is a convergent interval-function whose kernel is equal to $f(x)$.

The proof just given together with the proof that the limit point-function of a convergent interval-function is the limit of a sequence of continuous functions, and conversely, furnishes another proof of Baire's theorem.

The above theorems and proofs may readily be extended to the two-dimensional case, if we substitute interiors of squares for open intervals. An immediate consequence is the

THEOREM. *If a function $f(x, y)$, defined on an open region of the xy-plane, is the limit of a sequence of continuous functions, it has a point of continuity on every set SR with respect to SR, where S is a perfect set and R is an open region of the xy-plane.*

ST. MICHAEL'S COLLEGE.

THE EXTENSION OF RECTANGLE FUNCTIONS

By P. Reichelderfer and L. Ringenberg

Introduction

0.1. The theory of Lebesgue measure extends in a completely additive way the concept of area from a class of simple figures—circles, triangles, rectangles, and polygons in general—to a more comprehensive class of sets which is closed under the operations of addition, subtraction, multiplication, and limit. It has inspired many papers and books[1] dealing with the problem of securing a completely additive extension[2] of a general set function defined on a certain class of simple figures. For most applications, however, each of these general theories possesses certain artificial features, some of which we shall mention later (cf. §0.3). The purpose of this paper is to set forth a simple and natural theory which meets the needs of many applications. We shall restrict our considerations to the plane, since it at once offers rather "heavy" sets which may carry no weight and rather "meager" sets which may carry considerable weight. Before stating our results, we shall summarize certain important results in the literature. Results of Radon (cf. §0.2) will be carefully summarized since they will be used in establishing our theorems. Results of Caccioppoli (cf. §0.4) will also be stated because, while they would be quite useful in the applications, unfortunately they are false; we shall give a counter example to show this in §0.4.

0.2. Let R_0 be a fixed rectangle in the xy-plane bounded by the lines $x = 0$, $x = x_0 > 0$; $y = 0$, $y = y_0 > 0$.[3] A class of sets in R_0 is said to be *closed* (relative to R_0)[4]—and is denoted generically by K—if the following conditions are satisfied:

(i) every open set (relative to R_0) is in K;

(ii) if e is a set in K, then $R_0 - e$ is also a set in K;

Received July 10, 1940.

[1] See the bibliography at the end of this paper. Numbers in square brackets refer to references in this bibliography.

[2] For definitions of such concepts as set functions, completely additive set functions, completely additive extensions, etc., see, for example, [3], [4], [6].

[3] Rectangles, as R_0, whose sides are parallel by pairs to the x- and y-axes respectively are termed *oriented* rectangles.

[4] Radon (cf. [3]) and de la Vallée-Poussin (cf. [6]) define "Classe T" and "corps fermé" respectively; for sets in R_0, each of these concepts is equivalent to our closed class.

(iii) if $e_1, e_2, \cdots$ is a finite or denumerable sequence of sets in K, then $\sum_n e_n$ is also a set in K.[5]

Clearly every closed class K contains all Borel sets in R_0.[6]

Let $[\Phi, K]$ denote a completely additive set function defined on a closed range K in R_0. Radon [3] has shown that there always exist two non-negative completely additive set functions Φ_1 and Φ_2 defined on the same range as Φ and such that $\Phi(E) = \Phi_1(E) - \Phi_2(E)$ for every set E in K. Thus in investigating the question as to whether a given set function, defined on a certain class of sets in R_0, admits of a completely additive extension to a closed range, it is legitimate to express the given set function as the difference of two non-negative set functions, then to consider the question for these non-negative set functions.

Given a non-negative completely additive set function $[\Phi, K]$, Radon (cf. [3]) associates with it a point function $f(x, y)$ defined in the closed rectangle R_0 by the relations[7]

$$(1)\quad f(x, y) = \begin{cases} 0 & \text{if either } \begin{cases} 0 = x \\ 0 \leqq y \leqq y_0 \end{cases} \text{ or } \begin{cases} 0 \leqq x \leqq x_0 \\ 0 = y \end{cases}; \\ \Phi([0, x; 0, y)) & \text{if } \begin{cases} 0 < x \leqq x_0 \\ 0 < y \leqq y_0 \end{cases}. \end{cases}$$

Radon observes that for all values of x', x'', y', y'' satisfying the relations $0 \leqq x' < x'' \leqq x_0$, $0 \leqq y' < y'' \leqq y_0$ it is true that

$$(2)\qquad \Phi([x', x''; y', y'')) = f(x'', y'') - f(x'', y') - f(x', y'') + f(x', y').$$

Thus the function $f(x, y)$ has the property that, for all values of x', x'', y', y'' satisfying $0 \leqq x' \leqq x'' \leqq x_0$, $0 \leqq y' \leqq y'' \leqq y_0$, it is true that

$$(3)\qquad f(x'', y'') - f(x'', y') - f(x', y'') + f(x', y') \geqq 0.$$

Radon further shows that $f(x, y)$ has the property that, for every x', y' satisfying $0 < x' \leqq x_0$, $0 < y' \leqq y_0$, it is true that

$$(4)\qquad \lim f(x' - h, y' - k) = f(x', y')$$

when the non-negative variables h and k converge to zero in an arbitrary way. Next, *Radon exhibits, for every function $f(x, y)$ defined in the closed rectangle R_0 vanishing for $x = 0$ and for $y = 0$ and satisfying conditions* (3) *and* (4), *a non-*

[5] The reader will easily verify that (iv) the empty set is in K; (v) if $e_1, e_2, \cdots$ is a finite or denumerable sequence of sets in K, then $\prod_n e_n$ is also a set in K; (vi) if e_1 and e_2 are two sets in K, then $e_1 - e_2$ is also a set in K.

[6] In fact, the product of any number of closed classes is again a closed class. The class of all Borel sets in R_0 is identical with the product of all closed classes in R_0.

[7] For two points (x', y'), (x'', y'') subject to $x' < x''$, $y' < y''$ the notation $[x', x''; y', y'')$ represents the semi-open rectangle consisting of all points (x, y) satisfying $x' \leqq x < x''$, $y' \leqq y < y''$.

negative completely additive set function $[\Phi_0, K_0]$ *defined on a closed range* K_0, *bearing the relation* (2) *to* $f(x, y)$ *and also possessing the following property*:

(*) *Given* $\epsilon > 0$, *every set* $E \epsilon K_0$ *contains a closed set* F *such that* $\Phi_0(E) - \Phi_0(F) < \epsilon$.

Moreover, Radon observes that this extension is unique in the following sense: Let $[\Phi, K]$ be any non-negative completely additive set function defined on a closed range K; if $f(x, y)$ is the point function associated with $[\Phi, K]$ by the relation (1), then to $f(x, y)$ there belongs the non-negative completely additive set function $[\Phi_0, K_0]$ exhibited by Radon. On every set E which is in both K and K_0, it is true that $\Phi(E) = \Phi_0(E)$. Further, if K possesses property (*), then $K \subset K_0$; that is, $[\Phi_0, K_0]$ is *maximal* with respect to property (*). Finally, Radon has a similar result for arbitrary set functions.[8]

0.3. Radon (cf. §0.2) uses semi-open rectangles because it is possible to build open sets from them in an additive way. His theory is applicable directly to set functions defined on semi-open rectangles in R_0, and necessitates the consideration of an auxiliary point function. But when a function of rectangles arises in the applications, no sides play any special rôle—the value of the function belongs either to the open rectangle or to the closed rectangle. Thus it is quite desirable to have necessary and sufficient conditions in order that a set function defined on a class of open rectangles, or of closed rectangles, admit a completely additive extension.

0.4. Caccioppoli (cf. [1]) states the following theorem: Let $[\phi, \mathcal{F}]$ be a set function ϕ defined on the class $\mathcal{F}$ of all closed sets in the closed rectangle R_0 and possessing the following properties:

(i) $\phi(F) \geqq 0$ for every set $F \epsilon \mathcal{F}$;

(ii) there exists a positive number M such that $\phi(F) \leqq M$ for every set $F \epsilon \mathcal{F}$;

(iii) if F_1 and F_2 are two mutually exclusive sets in $\mathcal{F}$, then $\phi(F_1 + F_2) = \phi(F_1) + \phi(F_2)$;

(iv) if F_1 and F_2 are two sets in $\mathcal{F}$ such that $F_1 \subset F_2$, then $\phi(F_1) \leqq \phi(F_2)$;

(v) if F_ϵ denotes the (closed) set of points in R_0 whose distance from F does not exceed ϵ, where F is any set in $\mathcal{F}$, then $\lim \phi(F_\epsilon) = \phi(F)$ as ϵ tends to zero.

Then $[\phi, \mathcal{F}]$ possesses a completely additive extension to the class of all Borel sets in R_0.

This would be a very useful set of conditions were it not for the fact that there exist set functions $[\phi, \mathcal{F}]$ satisfying these five conditions, yet possessing no completely additive extension to the class of all Borel sets in R_0.[9] Consider the following example: Let S be a fixed closed square in the interior of R_0; define

[8] De la Vallée-Poussin (cf. [5], [6]) has sketched a theory for the extension of set functions which is similar in many respects to that of Radon. He replaces the semi-open rectangles used by Radon by a certain class of rectangles whose boundaries carry no weight for the function and which are, in a sense, everywhere dense in R_0.

[9] This fact was called to our attention by Professor Radó.

$\phi(F) = 1$ for $F \in \mathcal{F}$, provided $F \supset S$; define $\phi(F) = 0$ for $F \in \mathcal{F}$ otherwise. The reader will easily verify that the set function $[\phi, \mathcal{F}]$ satisfies all five conditions proposed by Caccioppoli; in fact, $[\phi, \mathcal{F}]$ is a completely additive set function. Yet $[\phi, \mathcal{F}]$ possesses no completely additive extension to the class of all Borel sets in R_0. For suppose there did exist one; denote it by $[\Phi_0, K_0]$, where K_0 is now the class of all Borel sets in R_0. Let F_1 and F_2 be two closed sets in $\mathcal{F}$ such that $F_1 + F_2$ contains S, but neither F_1 nor F_2 contains S. Using a well-known identity for additive set functions, one easily verifies that $1 = \phi(F_1 + F_2) + \phi(F_1 \cdot F_2) = \Phi_0(F_1 + F_2) + \Phi_0(F_1 \cdot F_2) = \Phi_0(F_1) + \Phi_0(F_2) = \phi(F_1) + \phi(F_2) = 0$ and this is false.[10]

0.5. In this paper we establish the following four theorems, all useful in the applications.

$$\text{THEOREM} \left\{\begin{matrix} 1. \\ 2. \\ 3. \\ 4. \end{matrix}\right. \quad \textit{Let} \left\{\begin{matrix} C^0 \\ C_0^0 \\ C \\ C_0 \end{matrix}\right. \textit{denote the class of all} \left\{\begin{matrix} \textit{oriented}^{11}\ \textit{open} \\ \textit{open} \\ \textit{oriented closed} \\ \textit{closed} \end{matrix}\right. \textit{rectangles in the}$$

$$\textit{fixed oriented rectangle } R_0 \text{ (cf. §0.2)}, \textit{ regarded as} \left\{\begin{matrix} \textit{open} \\ \textit{open} \\ \textit{closed} \\ \textit{closed} \end{matrix}\right. . \quad \textit{A necessary and suffi-}$$

$$\textit{cient condition that a given set function} \left\{\begin{matrix} [\phi, C^0] \\ [\phi, C_0^0] \\ [\phi, C] \\ [\phi, C_0] \end{matrix}\right. \textit{admit a completely additive exten-}$$

$$\textit{sion to a closed range in } R_0 \textit{ is that it satisfy condition} \left\{\begin{matrix} \mathfrak{C}^0\colon \textit{If } r_1^0, r_2^0, \cdots{}^{12} \\ \mathfrak{C}_0^0\colon \textit{If } r_1^0, r_2^0, \cdots \\ \mathfrak{C}\ \colon \textit{If } r_1, r_2, \cdots \\ \mathfrak{C}_0\colon \textit{If } r_1, r_2, \cdots \end{matrix}\right. \textit{is a finite}$$

$$\textit{or denumerable sequence of mutually exclusive rectangles in} \left\{\begin{matrix} C^0 \\ C_0^0 \\ C \\ C_0 \end{matrix}\right. , \textit{ and if}$$

$$\left\{\begin{matrix} R_1^0, R_2^0, \cdots \\ R_1^0, R_2^0, \cdots \\ R_1, R_2, \cdots \\ R_1, R_2, \cdots \end{matrix}\right. \textit{is any finite or denumerable sequence of rectangles in} \left\{\begin{matrix} C^0 \\ C_0^0 \\ C \\ C_0 \end{matrix}\right. \textit{such that}$$

[10] Caccioppoli (cf. [1]) has an analogous theorem for a function of open sets. An example similar to the one given above shows that this theorem is also false.

[11] See footnote 3.

[12] The symbols r, or R, denote generically rectangles which are closed relative to R_0; r^0 and R^0 denote the interior of r and R respectively.

$$\left\{\begin{array}{l} \sum\limits_m R_m^0 \supset \sum\limits_n r_n^0 \\ \sum\limits_m R_m^0 \supset \sum\limits_n r_n^0 \\ \sum\limits_m R_m \supset \sum\limits_n r_n \\ \sum\limits_m R_m \supset \sum\limits_n r_n \end{array}\right. , \textit{ then } \left\{\begin{array}{l} \sum\limits_n \phi(r_n^0) \leqq \sum\limits_m \phi(R_m^0) \\ \sum\limits_n \phi(r_n^0) \leqq \sum\limits_m \phi(R_m^0) \\ \sum\limits_n \phi(r_n) \leqq \sum\limits_m \phi(R_m) \\ \sum\limits_n \phi(r_n) \leqq \sum\limits_m \phi(R_m) \end{array}\right. .$$ [13]

These theorems may be proved by direct methods analogous to those used in the theory of Lebesgue measure. However, we shall derive these theorems from those of Radon quoted in §0.3; then theorems on uniqueness and on rectangle functions of variable sign are ready consequences of his theory. Part 1 contains formal verification of Theorems 1 and 2; Part 2 sketches proofs for Theorems 3 and 4.[14]

1. Extension of non-negative set functions of open rectangles

1.1. The fixed oriented rectangle R_0 (cf. §0.2), regarded as open, and a set function $[\phi, C^0]$ satisfying $\mathfrak{C}^0$ are given (cf. §0.5).

A line segment is called an *oriented* line, and is denoted generically by l, if it is the product of R_0 and of a line parallel to one of the axes. Given an oriented line l, if $R_1^0, R_2^0, \cdots$[15] is any infinite sequence of oriented rectangles in R_0 satisfying $R_n^0 \supset l$ for every n and $\lim_n | R_n | = 0$,[16] then, since $[\phi, C^0]$ satisfies $\mathfrak{C}^0$, it follows that $\phi(R_n^0)$ converges. Denote this limit by $\phi(l)$. The reader will verify that the number of oriented lines l for which $\phi(l) > \alpha$, where α is a positive number, is finite. Hence oriented lines l for which $\phi(l) = 0$ are dense both horizontally and vertically. Such lines will be called *regular* lines. Oriented rectangles whose boundaries are formed by segments of regular lines are called *regular* rectangles. The following lemma is an immediate consequence of these definitions and of $\mathfrak{C}^0$.

LEMMA. *If R^0 is a regular rectangle in C^0 and $\epsilon > 0$ is given, then there exists a regular rectangle $R_\epsilon^0 \epsilon C^0$ such that $R_\epsilon^0 \supset R$ and $\phi(R_\epsilon^0) < \phi(R^0) + \epsilon$.*

1.2. If R is any closed rectangle, then any set consisting of a finite number of non-overlapping rectangles $R_1, \cdots, R_n$ such that $R = \sum_{j=1}^{n} R_j$ is called a *subdivision* of R, and is denoted by $D(R)$. The *norm* of a subdivision $D(R)$, denoted by $\| D(R) \|$, is the maximum of the diameters of the elements in $D(R)$. The line segments which form the boundaries of the elements in $D(R)$ and are

[13] This condition was suggested by Professor Radó, whom we wish to thank for access to unpublished materials and for advice in the preparation of this paper.

[14] We prove in this paper that the conditions which are stated are sufficient conditions. The necessity of the condition may be shown in each instance by a simple proof.

[15] See footnote 12.

[16] If E is any set, then $| E |$ denotes the exterior measure of E.

not part of the boundary of R are called the *lines of subdivision* of $D(R)$. A subdivision $D(R)$ is termed *elementary* if all the lines of subdivision of $D(R)$ extend from boundary to boundary of R. A sequence $D_1(R)$, $D_2(R)$, $\cdots$ of subdivisions of R is termed *nested* if all the lines of subdivision of $D_j(R)$ are also lines of subdivision of $D_{j+1}(R)$ for $j = 1, 2, \cdots$.

A subdivision of any oriented open rectangle is termed *regular* if all the lines of subdivision are segments of regular lines (cf. §1.1). Every rectangle in C^0 possesses regular subdivisions—in fact elementary regular subdivisions—with arbitrarily small norms, since the regular lines are dense (cf. §1.1). Clearly a regular subdivision of a regular rectangle consists of a finite number of regular rectangles. As an easy consequence of $\mathfrak{C}^0$ and of the lemma in §1.1 we have the

LEMMA. *If R^0 is a regular rectangle in C^0, if $D(R)$ is a regular subdivision of R, then $\phi(R^0) = \sum \phi(r^0)$ for $r \in D(R)$.*

1.3. In contrast to the lemma in §1.1 we have the following

LEMMA. *If R^0 is any rectangle in C^0 and $\epsilon > 0$ is given, then there exists a regular rectangle $R^0_\epsilon \in C^0$ such that $R_\epsilon \subset R^0$ and $\phi(R^0_\epsilon) > \phi(R^0) - \epsilon$.*

Proof. Let $D_1(R)$, $D_2(R)$, $\cdots$ be a nested sequence of regular elementary subdivisions of R for which $\lim_n \| D_n(R) \| = 0$ (cf. §1.2). Define the following classes of rectangles:

K_1 : class of all r such that $r \in D_1(R)$, $r \cdot b(R) = 0$.[17]

K_n : class of all r such that either $r \in K_{n-1}$; or $r \in D_n(R)$, r is not contained in any element of K_{n-1}, and $r \cdot b(R) = 0$.

K: class of all r such that $r \in K_n$ for some n.

Every class K_n consists of a finite number of regular rectangles the sum of which is a regular rectangle in C^0. Denote it by R^0_n. Given $\epsilon > 0$, it is easily seen that there exists an $n = n(\epsilon)$ such that $\phi(R^0) - \epsilon < \sum \phi(r^0)$ for $r \in K_n$. Then, since $\phi(R^0_n) = \sum \phi(r^0)$ for $r^0 \in K_n$ by the lemma in §1.2, it follows that $\phi(R^0) - \epsilon < \phi(R^0_n)$; thus R^0_n qualifies as the rectangle R^0_ϵ in the lemma.

1.4. We now introduce some notation which will simplify the proofs of several lemmas in the sequel. An oriented rectangle bounded by the lines $x = x'$, $x = x''$, $y = y'$, $y = y''$, where $x' < x''$, $y' < y''$, will be denoted by $(x', x''; y', y'')$ if it is regarded as open (in the absolute sense), by $[x', x''; y', y'']$ if it is regarded as closed (in the absolute sense), and by $[x', x''; y', y'')$, or by r', or by R', if it is regarded as semi-open (cf. footnote 7). Let $R = [x', x''; y', y'']$ be any oriented rectangle in R_0 for which $x' > 0$, $y' > 0$ (cf. §0.2). Let

$$R_1 = [0, x'; 0, y'], \qquad R_{12} = [0, x'; 0, y''],$$

$$R_{13} = [0, x''; 0, y'], \qquad R_{123} = [0, x''; 0, y''].$$

[17] If E is any set, then $b(E)$ denotes the boundary of E.

Given $\epsilon > 0$, it follows by the lemma in §1.3 and by $\mathfrak{C}^0$ that there exists an $\eta = \eta(\epsilon)$ such that $R_{12\epsilon} = [0, x' - \eta; 0, y'' - \eta]$ and $R_{13\epsilon} = [0, x'' - \eta; 0, y' - \eta]$ are regular rectangles satisfying $\phi(R^0_{12\epsilon}) > \phi(R^0_{12}) - \epsilon$, $\phi(R^0_{13\epsilon}) > \phi(R^0_{13}) - \epsilon$. Let

$$R_{1\epsilon} = [0, x' - \eta; 0, y' - \eta], \qquad R_{123\epsilon} = [0, x'' - \eta; 0, y'' - \eta],$$
$$R_{4\epsilon} = [x' - \eta, x'' - \eta; y' - \eta, y'' - \eta].$$

1.5. Consider now the point function

$$f(x, y) = \begin{cases} 0 & \text{if either} \begin{cases} 0 \leqq x \leqq x_0 \\ 0 = y \end{cases} \text{or} \begin{cases} 0 = x \\ 0 \leqq y \leqq y_0 \end{cases}, \\ \phi((0, x; 0, y)) & \text{if} \begin{cases} 0 < x \leqq x_0 \\ 0 < y \leqq y_0 \end{cases}. \end{cases}$$

Clearly $f(x, y)$ is real, finite, and single-valued on the closed rectangle R_0. Set $F(x', x''; y', y'') = f(x'', y'') - f(x'', y') - f(x', y'') + f(x', y')$ for $0 \leqq x' \leqq x'' \leqq x_0$, $0 \leqq y' \leqq y'' \leqq y_0$.

LEMMA 1. *$F(x', x''; y', y'') \geqq 0$ for all x', x'', y', y'' satisfying $0 \leqq x' \leqq x'' \leqq x_0$, $0 \leqq y' \leqq y'' \leqq y_0$.*

Proof. We shall discuss only the case where $0 < x' < x''$, $0 < y' < y''$, leaving the discussion of the other cases for the reader. Using the notation in §1.4, the lemma in §1.2, and $\mathfrak{C}^0$, we verify, for every $\epsilon > 0$, that

$$\begin{aligned} F(x', x''; y', y'') &= \phi(R^0_{123}) - \phi(R^0_{12}) - \phi(R^0_{13}) + \phi(R^0_1) \\ &\geqq \phi(R^0_{123\epsilon}) - \phi(R^0_{12\epsilon}) - \phi(R^0_{13\epsilon}) + \phi(R^0_{1\epsilon}) - 2\epsilon \\ &\geqq \phi(R^0_{4\epsilon}) - 2\epsilon \geqq -2\epsilon. \end{aligned}$$

Thus the lemma is established in this case.

LEMMA 2. *If (x', y') is any point such that $0 < x' \leqq x_0$, $0 < y' \leqq y_0$, then $f(x', y') = \lim f(x' - h, y' - k)$ as the non-negative variables h and k approach zero.*

This lemma follows immediately from the lemma in §1.3 and $\mathfrak{C}^0$.

1.6. In view of Lemmas 1 and 2 in §1.5 it is clear that $f(x, y)$ satisfies conditions (3) and (4) of Radon (cf. §0.2); thus there exists a non-negative completely additive set function $[\Phi_0, K_0]$ defined on a closed range K_0 and such that $\Phi_0([x', x''; y', y'']) = F(x', x''; y', y'')$. We shall now show that $[\Phi_0, K_0]$ is an extension of our original set function $[\phi, C^0]$ (cf. §1.1).

1.7. LEMMA 1. *If R^0 is a regular rectangle in C^0, then $\phi(R^0) = \Phi_0(R^0)$.*

Proof. We shall discuss the case when $b(R) \subset R_0$, leaving other cases for the reader. Using the lemma in §1.2 and the notation of §1.5, we have $\Phi_0(R') =$

$F(x', x''; y', y'') = \phi(R^0)$. Since Φ_0 is non-negative and completely additive, it follows that $\Phi_0(R') \geqq \Phi_0(R^0)$. But, given $\epsilon > 0$, there exists a regular rectangle $R^0_\epsilon \epsilon C^0$ such that $R_\epsilon \subset R^0$, and $\phi(R^0_\epsilon) > \phi(R^0) - \epsilon$ (cf. §1.3). Thus by the preceding relations,

$$\phi(R^0) - \epsilon < \phi(R^0_\epsilon) = \Phi_0(R'_\epsilon) \leqq \Phi_0(R^0) \leqq \Phi_0(R') = \phi(R^0).$$

But ϵ is arbitrary, thus $\phi(R^0) = \Phi_0(R^0)$.

LEMMA 2. *If R^0 is any rectangle in C^0, then $\phi(R^0) = \Phi_0(R^0)$.*

Proof. It follows from the lemma in §1.3 that there exists a sequence of regular rectangles $R^0_n \epsilon C^0$ such that $R_n \subset R^0$, $\phi(R^0) \geqq \phi(R^0_n) > \phi(R^0) - n^{-1}$, and $R^0 - R^0_n$ approaches the null set as $n \to \infty$. In view of Lemma 1 and the properties of Φ_0, it follows that

$$\begin{aligned}\phi(R^0) &= \lim_n [\phi(R^0) - n^{-1}] = \lim_n \phi(R^0_n) = \lim_n \Phi_0(R^0_n) \\ &= \lim_n [\Phi_0(R^0) - \Phi_0(R^0 - R^0_n)] = \Phi_0(R^0).\end{aligned}$$

Thus the lemma is established.

1.8. The lemmas in §1.7 show that $[\Phi_0, K_0]$ is a non-negative completely additive extension of the set function $[\phi, C^0]$ to a closed range K_0. Thus the first theorem stated in §0.5 is established.

1.9. We now proceed to prove the second theorem. R^0 is the fixed oriented open rectangle (cf. §0.2), and a non-negative set function $[\phi, C^0_0]$ satisfying condition $\mathfrak{C}^0_0$ is given (cf. §0.5). It is clear that this function ϕ defined on the subclass C^0 is a set function satisfying $\mathfrak{C}^0$, and hence possessing a non-negative completely additive extension $[\Phi_0, K_0]$ to a closed range (cf. §1.8). We now show that $[\Phi_0, K_0]$ is a non-negative completely additive extension of $[\phi, C^0_0]$ (cf. §0.5). This follows immediately from the

LEMMA. *If R^0 is any rectangle in C^0_0, then $\phi(R^0) = \Phi_0(R^0)$.*

Proof. The rectangle R^0 may be expressed as the sum of a denumerable number of mutually exclusive, semi-open, regular rectangles $r'_1, r'_2, \cdots$. Since Φ_0 is additive, we have by §1.7

$$\Phi_0(R^0) = \sum_i \Phi_0(r'_i) = \sum_i \phi(r^0_i).$$

Given $\epsilon > 0$, since $r^0_i \epsilon C^0$ is regular, there exists a regular rectangle $r^0_{i\epsilon} \epsilon C^0$ such that $r^0_{i\epsilon} \supset r_i$ and $\phi(r^0_{i\epsilon}) < \phi(r^0_i) + 2^{-i}\epsilon$ for $i = 1, 2, \cdots$ (cf. §1.1). It now follows from $\mathfrak{C}^0_0$ and the preceding relations that

$$\phi(R^0) \geqq \sum_i \phi(r^0_i) = \Phi_0(R^0) > \sum_i \phi(r^0_{i\epsilon}) - \epsilon \geqq \phi(R^0) - \epsilon.$$

Since ϵ is arbitrary, it follows that $\phi(R^0) = \Phi_0(R^0)$.

2. Extension of non-negative set functions of closed rectangles

2.1. In this part R_0 (cf. §0.2) will denote a closed rectangle and C will denote the class of all oriented closed rectangles in R_0. Given a non-negative set function $[\phi, C]$ satisfying condition $\mathfrak{C}$ (cf. §0.5), we shall show that $[\phi, C]$ possesses a non-negative completely additive extension to a closed range (cf. §0.2), again using the results of Radon (cf. §0.2). The procedure will be similar to that used in Part 1 for the open rectangle functions; we shall give details on points where the theories differ and omit details at points which are similar to those for the open rectangle functions considered in Part 1.

2.2. The definition of an oriented line is analogous to that in §1.1, but here R_0 is a closed rectangle (cf. §2.1). Given an oriented line l, define $\phi(l) =$ l.u.b. $\limsup_n \phi(R_n)$ for all sequences $R_1, R_2, \cdots$ of oriented rectangles in R_0 satisfying $R_n \supset l$ and $\lim_n |R_n| = 0$.[18] The definitions and properties of regular lines, and regular rectangles and regular subdivisions are analogous to those given in §§1.1 and 1.2. The following lemmas which are the analogues of the lemmas in §§1.1 and 1.2 respectively are immediate consequences of $\mathfrak{C}$ and of these definitions:

LEMMA 1. *If R is any regular rectangle in C, and $\epsilon > 0$ is given, then there exists a regular rectangle $R_\epsilon \in C$ such that $R_\epsilon \subset R^0$ and $\phi(R_\epsilon) > \phi(R) - \epsilon$.*

LEMMA 2. *If R is any regular rectangle in C, if $D(R)$ is any regular subdivision of R, then $\phi(R) = \sum \phi(r)$ for $r \in D(R)$.*

2.3. Our discussion is now divided into cases. First, we shall assume in §§2.4–2.6 that R_0 is a regular rectangle; that is, $\phi(l) = 0$ for $l \subset b(R_0)$.[19] In §§2.7–2.10 we shall treat the general case.

2.4. Analogous to the lemma in §1.3 we have the

LEMMA. *If we assume that R_0 is regular, if R is any rectangle in C, and $\epsilon > 0$ is given, then there exists a regular rectangle R_ϵ such that $R_\epsilon^0 \supset R$ and $\phi(R_\epsilon) < \phi(R) + \epsilon$.*[20]

Proof. Let $D_1(R_0), D_2(R_0), \cdots$ be a nested sequence of regular elementary subdivisions of R_0 for which $\lim_n \| D_n(R_0) \| = 0$ (cf. §1.2). Define the following classes of rectangles:

K_1 : class of all r such that $r \in D_1(R_0)$, $r \cdot R = 0$.

[18] The definition of $\phi(l)$ in §1.1 is equivalent to the analogue of the definition given here, i.e., to the definition obtained by replacing closed rectangles by open rectangles in this definition.

[19] See footnote 17.

[20] This lemma is obviously false if R_0 is not a regular rectangle.

K_n : class of all r such that either $r \epsilon K_{n-1}$; or $r \epsilon D_n(R_0)$, r is not contained in any element of K_{n-1}, and $r \cdot R = 0$.

K: class of all r such that $r \epsilon K_n$ for some n.

Clearly the rectangles in $K - K_n$, together with R, form a regular rectangle containing R in its interior. Denote it by R_n. It follows easily from the lemmas in §2.2 and from $\mathfrak{C}$ that $\phi(R_n) = \phi(R) + \sum \phi(r)$ for $r \epsilon K - K_n$ for each n. But, given $\epsilon > 0$, it is readily seen that there exists an $n = n(\epsilon)$ such that $\sum \phi(r)$ for $r \epsilon K - K_n$ is less than ϵ. Thus the lemma is established.

2.5. Let us now consider the point function (cf. §1.5):

$$f(x, y) = \begin{cases} 0 & \text{if either} \begin{cases} 0 \leqq x \leqq x_0 \\ 0 = y \end{cases} \text{or} \begin{cases} 0 = x \\ 0 \leqq y \leqq y_0 \end{cases}, \\ \lim\limits_{\substack{h > 0 \to 0 \\ k > 0 \to 0}} \phi([0, x - h; 0, y - k]) & \text{if} \begin{cases} 0 < x \leqq x_0 \\ 0 < y \leqq y_0 \end{cases}. \end{cases}$$

Clearly $f(x, y)$ is real, finite, and single-valued on the closed rectangle R_0. Set $F(x', x''; y', y'') = f(x'', y'') - f(x'', y') - f(x', y'') + f(x', y')$ for $0 \leqq x' \leqq x'' \leqq x_0$, $0 \leqq y' \leqq y'' \leqq y_0$.

LEMMA 1. *$F(x', x''; y', y'') \geqq 0$ for all x', x'', y', y'' satisfying $0 \leqq x' \leqq x'' \leqq x_0$, $0 \leqq y' \leqq y'' \leqq y_0$.*

The proof follows from the lemma in §2.2 by a reasoning similar to that in the proof of Lemma 1 in §1.5.

LEMMA 2. *If (x', y') is any point such that $0 < x' \leqq x_0$, $0 < y' \leqq y_0$, then $f(x', y') = \lim f(x' - h, y' - k)$, where h and k are arbitrary non-negative variables approaching zero.*

The proof follows immediately from the definition of $f(x, y)$ and from $\mathfrak{C}$.

2.6. From Lemmas 1 and 2 in §2.5 it is clear that $f(x, y)$ satisfies the conditions (3) and (4) of Radon (cf. §0.2); thus there exists a non-negative completely additive set function $[\Phi_0, K_0]$ defined on a closed range K_0 and such that $\Phi_0([x', x''; y', y'']) = F(x', x''; y', y'')$. A reasoning similar to that used in §1.7 shows that $[\Phi_0, K_0]$ is a non-negative completely additive extension of the set function $[\phi, C]$.

2.7. Now we turn to the general case where R_0 is any rectangle, not necessarily regular.

LEMMA. *Given $\epsilon > 0$, there exist a regular rectangle $R \epsilon C$ and a sequence of regular rectangles $r_{j\epsilon} \epsilon C$ satisfying $R \cdot b(R_0) = 0$,*

$$\sum_j r_{j\epsilon} \supset R_0^0 - R, \qquad \sum_j \phi(r_{j\epsilon}) < \epsilon.$$

Proof. Let $D_1(R_0), D_2(R_0), \cdots$ be a nested sequence of regular elementary subdivisions of R_0 for which $\lim_n \| D_n(R_0) \| = 0$ (cf. §1.2). Using the notations of §1.3 with R replaced by R_0, Lemma 2 in §2.2, and condition $\mathfrak{C}$, we have, for every n, $\phi(R_0) \geqq \phi(R_n) = \sum \phi(r)$ for $r \in K_n$. These relations clearly imply the lemma.

2.8. Let R_* be a closed oriented rectangle containing R_0 in its interior; let C_* denote the class of all closed oriented rectangles contained in R_*. Define

$$\phi_*(R) = \begin{cases} \phi(R \cdot R_0) & \text{if } R_0 \cdot R \in C,\ R \in C_*; \\ 0 & \text{if } R \cdot R_0 \text{ is empty, } R \in C_*; \\ \text{g.l.b. } \phi(r) & \text{for all } r \in C \text{ such that } R \cdot R_0 \subset r \subset R_0, \text{ if } R \cdot R_0 \text{ is a line} \\ & \quad \text{segment or a point, } R \in C_*. \end{cases}$$

Clearly $[\phi_*, C_*]$ is a non-negative function defined on all closed oriented rectangles in R_* with respect to which R_* is a regular rectangle. Denote by $\mathfrak{C}_*$ the condition for $[\phi_*, C_*]$ on R_* which is analogous to condition $\mathfrak{C}$ for $[\phi, C]$ on R_0 (cf. §0.5). Then $[\phi_*, C_*]$ satisfies condition $\mathfrak{C}_*$; this[21] follows from the

LEMMA. *If $r_1, r_2, \cdots, r_n$ is a finite set of mutually exclusive rectangles in C_* and $R_1, R_2, \cdots$ is a finite or denumerable sequence of rectangles in C_* such that $\sum_m R_m \supset \sum_{k=1}^n r_k$, then $\sum_m \phi_*(R_m) \geqq \sum_{k=1}^n \phi_*(r_k)$.*

Proof. Given $\epsilon > 0$ there exist a regular rectangle $R \in C$ and a sequence of regular rectangles $r_{j\epsilon} \in C$ satisfying $R \cdot b(R_0) = 0$, $\sum_j r_{j\epsilon} \supset R_0^0 - R$, $\sum_j \phi(r_{j\epsilon}) < \epsilon$ (cf. §2.7). It is clear from the definition of ϕ_* that we may replace the given sequence of rectangles R_m by a sequence of rectangles R_m'' satisfying the following conditions:

(i) $R_m'' = R_m$ if $R_m \cdot R_0$ is a rectangle in C or is empty;

(ii) $R_m'' \supset R_m$, $R_m'' \cdot R_0$ is a rectangle in C and $\phi(R_m'' \cdot R_0) < \phi_*(R_m) + 2^{-m}\epsilon$ if $R_m \cdot R_0$ is a line segment or a point.

Next, we can replace the finite sequence of mutually exclusive rectangles r_k by a finite sequence of mutually exclusive rectangles $r_k'' \in C_*$ satisfying the following:

(i) $r_k'' = r_k$ if $r_k \cdot R_0$ is a rectangle in C or is empty;

(ii) $r_k'' \supset r_k$, $r_k'' \cdot R_0$ is a rectangle in C, and $r_k'' \cdot R_0 \subset \sum_j r_{j\epsilon} + \sum_m R_m'' \cdot R_0$ if $r_k \cdot R_0$ is a line segment or a point.

[21] Condition $\mathfrak{C}$ is obviously equivalent to the following condition: If $r_1, \cdots, r_n$ is any finite sequence of mutually exclusive closed rectangles in R_0, if $R_1, \cdots, R_m, \cdots$ is any finite or denumerable sequence of rectangles in R_0 such that $\sum_{k=1}^n r_k \subset \sum_m R_m$, then $\sum_{k=1}^n \phi(r_k) \leqq \sum_m \phi(R_m)$. Similarly, it is clear that $\mathfrak{C}_*$ is equivalent to the condition stated in the lemma.

Now clearly $r_1'' \cdot R_0, r_2'' \cdot R_0, \cdots, r_m'' \cdot R_0$ is a finite set of mutually exclusive rectangles in C, and $R_1'' \cdot R_0, r_{1\epsilon}, R_2'' \cdot R_0, r_{2\epsilon}, \cdots$ is a sequence of rectangles in C such that $\sum_m R_m'' \cdot R_0 + \sum_j r_{j\epsilon} \supset \sum_{k=1}^{n} r_k'' \cdot R_0$. From condition $\mathfrak{C}$ and the preceding relations we have

$$\sum_{k=1}^{n} \phi_*(r_k) \leqq \sum_{k=1}^{n} \phi_*(r_k'') = \sum_{k=1}^{n} \phi(r_k'' \cdot R_0)$$
$$\leqq \sum_m \phi(R_m'' \cdot R_0) + \sum_j \phi(r_{j\epsilon}) \leqq \sum_m \phi_*(R_m) + 2\epsilon.$$

Since $\epsilon > 0$ is arbitrary, the lemma follows.

2.9. Since $[\phi_*, C_*]$ is a non-negative function of closed oriented rectangles satisfying condition $\mathfrak{C}_*$ with respect to which R_* is regular, it follows by §2.6 that there exists a non-negative completely additive extension $[\Phi_*, K_*]$ of $[\phi_*, C_*]$ to a closed range K_* of sets in R_*. Denote by K_0 the class of all sets in K_* which are also in R_0; K_0 is a closed range of sets in R_0 (cf. footnote 5). It is obvious that $[\Phi_*, K_0]$ is a non-negative completely additive extension of $[\phi, C]$ to a closed range.

2.10. Again, let R_0 denote a closed rectangle, and denote by C_0 the class of all closed rectangles in R_0, whether oriented or not (cf. §0.5). If $[\phi, C_0]$ is any non-negative function of rectangles satisfying condition $\mathfrak{C}_0$ (cf. §0.5), then the associated non-negative function of rectangles $[\phi, C]$ satisfies condition $\mathfrak{C}$ (cf. §0.5), hence admits a non-negative completely additive extension $[\Phi_0, K_0]$ to a closed range. A reasoning similar to that in §1.9 shows that $[\Phi_0, K_0]$ is also a non-negative completely additive extension of $[\phi, C_0]$ to a closed range. Thus the results stated in §0.5 are established.

Bibliography

1. R. Caccioppoli, *Sulle funzioni additive d'insieme*, Rendiconti dell'Accademia delle Scienze Fisiche e Matematiche, Classe della Società Reale di Napoli, (3), vol. 34(1927), pp. 150–153.
2. A. J. Maria, *Functions of plurisegments*, Transactions of the American Mathematical Society, vol. 28(1926), pp. 448–471.
3. J. Radon, *Theorie und Anwendungen der absolut-additiven Mengenfunktionen*, Sitzungsberichte der mathematisch-naturwissenschaftlichen Klasse der Kaiserlichen Akademie der Wissenschaften, vol. 122(1913), pp. 1295–1438.
4. S. Saks, *Theory of the Integral*, Warsaw, 1937.
5. C. de la Vallée-Poussin, *Extension de la méthode du balayage de Poincaré et problème de Dirichlet, Correspondance entre les fonctions positives d'ensemble et les fonctions de point*, Annales de l'Institut Henri Poincaré, vol. 2(1932), pp. 169–232.
6. C. de la Vallée-Poussin, *Intégrales de Lebesgue, Fonctions d'Ensemble, Classes de Baire*, 2d ed., Paris, 1934, Chapter VI.

The Ohio State University.

KHINTCHINE'S PROBLEM IN METRIC DIOPHANTINE APPROXIMATION

By R. J. Duffin and A. C. Schaeffer

It has long been known that if x is any real number there exists an infinitude of rational numbers p/q which[1] satisfy

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^2}. \tag{1}$$

This relation raises the question: can the function $1/q^2$ on the right be replaced by a smaller function to obtain a sharper inequality? This question was answered by Hurwitz who showed that one can use the function $1/(\sqrt{5}q^2)$. Hurwitz' inequality is the "best possible" in the sense that if $1/\sqrt{5}$ is replaced by a smaller constant there are numbers x which can be approximated in the above manner only a finite number of times. An example is $x = \frac{1}{2}(1 + \sqrt{5})$.

It might be supposed, however, that by ignoring a certain limited class of numbers the inequality can be made sharper for the remaining numbers. A common method of ignoring exceptional sets is to use the idea of Lebesgue measure and to disregard sets of measure zero. The application of Lebesgue measure to improve inequality (1) was made by Khintchine[2] in 1924. However, other types of metrical problems in Diophantine approximation had been studied much earlier.

Khintchine's Theorem. *Let $\{\alpha_q\}$ be a sequence of positive numbers which satisfies*

(a) $\sum_{q=1}^{\infty} \alpha_q = \infty$,

(b) *$q\alpha_q$ is a decreasing function of q.*

Then for almost all x there exist arbitrarily many rational numbers p/q which satisfy

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q}. \tag{2}$$

An example of a sequence to which this theorem applies is $\alpha_q = (q \log q)^{-1}$.

One of the results of this paper is the replacement of Khintchine's condition (b) by the weaker condition

(b′) *α_q/q^c is a decreasing function of q.*

Here c may be any real constant.

Received August 12, 1940.

[1] The symbols p and q will be used throughout the text to denote positive integers.

[2] A. Khintchine, *Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen*, Mathematische Annalen, vol. 92(1924), pp. 115–125. *Zur metrischen Theorie der diophantischen Approximationen*, Mathematische Zeitschrift, vol. 24(1926), pp. 706–714.

It is important to notice the following distinction which now arises. If, for some x, Khintchine's inequality (2) is satisfied by integers p and q, then it is also satisfied for the same x by the relatively prime integers p_1 and q_1 obtained by dividing p and q by their greatest common divisor. This follows since, by (b), $\alpha_q/q \leqq \alpha_{q_1}/q_1$. A similar remark applies to inequality (1). With condition (b′) this is, however, no longer true since it may be that $\alpha_q/q > \alpha_{q_1}/q_1$. Thus we must distinguish between approximation by any integers p and q, and approximation by relatively prime integers. Our results suggest that the more natural formulation of this problem is in terms of reduced fractions.

The possibility of approximation by reduced fractions was noted by Walfisz. He showed that the following result[3] may be obtained from Khintchine's theorem.

WALFISZ' THEOREM. *Let $\{\alpha_q\}$ satisfy the conditions of Khintchine's theorem and the additional condition*:

(c) *α_q/α_{2q} is bounded.*

Then for almost all x there exist arbitrarily many relatively prime p and q such that

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q}$$

with $q \not\equiv 2 \pmod 4$.

Walfisz used his theorem to study the behavior of the function

$$1 + 2 \sum_{1}^{\infty} z^{n^2}$$

near the circle of convergence.

The above theorem will be a consequence of Theorem III of the present paper; in fact we find that condition (c) may be omitted without changing the conclusions of the theorem.

Condition (b′) may be replaced by the following condition, which we show is weaker than (b′),

(b″) *There is a constant $c > 0$ such that*

$$\sum_{\nu=1}^{n} \frac{\alpha_\nu \phi(\nu)}{\nu} > c \sum_{\nu=1}^{n} \alpha_\nu$$

for arbitrarily many n.

Here $\phi(\nu)$ is the Euler ϕ-function, $\phi(\nu)$ is the number of integers less than ν and relatively prime to ν.

In Theorem I we show that if the α_q are non-negative and satisfy conditions (a) and (b″) then the conclusions of Khintchine's theorem are still true. Under these conditions some of the α_q may be equal to zero, so our results show that the approximation

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q}$$

[3] A. Walfisz, *Ein metrischer Satz über Diophantische Approximationen*, Fundamenta Mathematicae, vol. 16(1930), pp. 361–385.

is possible even if the denominator q is required to belong to a suitable sequence of integers. On the other hand, the smoothness character of the α_q stated in (b) and (b′) implies that in Khintchine's theorem the denominators are required to range over all positive integers.

For example, let $\alpha_q = 0$ if q is composite and $\alpha_q = 1/q$ if q is prime. Condition (a) is satisfied since the sum of the reciprocals of the primes diverges. Moreover, since $\phi(q)/q \geqq \frac{1}{2}$ when q is prime, condition (b″) is satisfied. Thus we find that inequality (1) is satisfied in the metric sense even when q is required to be a prime.

As a second example, let $\alpha_q = k^{-1}$ if $q = 10^k$ $(k = 1, 2, 3, \cdots)$, and let $\alpha_q = 0$ otherwise. Then condition (a) is satisfied. Condition (b″) is likewise since $\phi(10^k)/10^k = \frac{2}{5}$. Thus for almost all x there exist arbitrarily many integers p and k such that[4]

$$| 10^k x - p | < k^{-1}.$$

While condition (b) in Khintchine's theorem has been weakened, it cannot be omitted altogether. This is shown by an example at the end of the present paper.

If inequality (2) is satisfied almost everywhere in (0, 1), then it will be satisfied almost everywhere in $(-\infty, \infty)$, so we need only consider the case $0 < x < 1$. Roughly speaking, the method is this: given the sequence $\{\alpha_q\}$, for each q let p/q, where p runs through all integers less than q and relatively prime to q, be the center of an open interval of length $2\alpha_q/q$. This defines for each q an open set E_q of measure of $2\alpha_q\phi(q)/q$. Let E be the sum of these sets, $E = \sum_1^\infty E_q$. If the measure of E is 1, then for almost all x inequality (2) will be satisfied for at least one pair of relatively prime p and q. In order to show that[5] $| E | = 1$ we assume conditions which insure that $\sum | E_q | = \sum 2\alpha_q\phi(q)/q$ diverges, and are led to the problem of estimating the overlapping of the sets E_q.

The principal theorem to be proved is

THEOREM I. *Let $\{\alpha_\nu\}$ be a sequence of non-negative numbers which satisfies the conditions*:

(a) $\sum_{\nu=1}^{\infty} \alpha_\nu = \infty$.

[4] Indeed the same method which is used in the proof of Theorem I will show that if conditions (a) and (b″) are satisfied then for almost all x there exist arbitrarily many relatively prime integers p and q such that $0 < x - p/q < \alpha_q/q$. A similar remark applies for the inequality $0 < p/q - x < \alpha_q/q$. Hence, in this example, we shall have

$$0 < 10^k x - p < 1/k.$$

Then if x is written in decimal form, for almost all x there will be arbitrarily large k such that the k-th digit is immediately followed by $[\log_{10} k]$ consecutive zeros.

[5] The symbol $| E |$ will be used to denote the measure of the set E.

(b″) *There is a constant* $c > 0$ *such that*

$$\sum_{\nu=1}^{n} \frac{\alpha_\nu \phi(\nu)}{\nu} > c \sum_{\nu=1}^{n} \alpha_\nu$$

for arbitrarily many n.
Then for almost all x *there exist arbitrarily many relatively prime* p *and* q *such that*

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q}.$$

The proof of Theorem I will depend on several lemmas.

LEMMA I. *Let* N *and* M *be given positive integers. The number of positive integer pairs* $\{x, y\}$ *which satisfy*

$$(3) \qquad 0 < | xN - yM | \leqq A,$$

where $1 \leqq x \leqq M$, $1 \leqq y \leqq N$, *is equal to or less than* $2A$.

This may be proved by a familiar argument of elementary number theory.

DEFINITION. If α satisfies $0 \leqq \alpha < \frac{1}{2}$ and $q > 1$, let E_q^α denote the set in $(0, 1)$ consisting of $\phi(q)$ open intervals each of length $2\alpha/q$ with centers at p/q, where p and q are relatively prime and $0 < p < q$.

The "probable" measure of the overlapping of two sets E_q^α and E_n^β would be $| E_q^\alpha | \, | E_n^\beta |$. The overlapping is in some cases more than this, but we do have the following.

LEMMA II. *If* $q \neq n$ *then*

$$(5) \qquad | E_q^\alpha E_n^\beta | \leqq 4\alpha\beta.$$

Proof. If an interval I' of E_q^α overlaps an interval I'' of E_n^β, then

$$0 < \left| \frac{p}{q} - \frac{m}{n} \right| < \frac{\alpha}{q} + \frac{\beta}{n}$$

or

$$0 < | np - mq | < n\alpha + q\beta,$$

where p/q is the center of I' and m/n is the center of I''. It is only a matter of notation to suppose that $\alpha/q \leqq \beta/n$. Then if I' and I'' have any points in common, we must have

$$0 < | np - mq | < 2q\beta.$$

By Lemma I, there are no more than $4q\beta$ solutions of this inequality. Hence the overlapping of the two sets E_n^β and E_q^α can be no more than

$$(4q\beta)(\alpha/q) = 4\alpha\beta.$$

In the proof of Theorem I we shall need an estimate of how regularly distributed are the numbers which are less than n and relatively prime to n.

LEMMA III. *If* $\phi_\lambda(n)$, $\lambda \geqq 0$, *denotes the number of positive integers which are equal to or less than* λn *and are relatively prime to* n, *then*

$$\phi_\lambda(n) = \phi(n)(\lambda + \rho),$$

where $|\rho| = |\rho(\lambda, n)| < An^{-\frac{1}{4}}$. *Here* A *is an absolute constant.*[6]

Proof. Let $n > 1$ be written in the canonical form

$$n = p_1^{\gamma_1} p_2^{\gamma_2} \cdots p_k^{\gamma_k}.$$

An exact expression for $\phi_\lambda(n)$ is

$$\phi_\lambda(n) = [\lambda n] - \sum [\lambda n / p_\nu] + \sum [\lambda n / p_\nu p_j] - \cdots , \tag{6}$$

where the primes occurring in each denominator are all different and are divisors of n. Here $[x]$ denotes the integral part of x. This may be proved by a familiar argument of elementary number theory.

If we remove the square brackets on the right side of (6), we obtain

$$\phi_\lambda(n) = \lambda\phi(n) + R,$$

where the error term R is less than the number of divisors of n, $d(n)$. But it is well known that

$$\frac{d(n)}{\phi(n)} < An^{-\frac{1}{4}},$$

so the lemma follows.

Suppose we are given a sequence $\{\alpha_n\}$ with $0 \leqq \alpha_n < \frac{1}{2}$. According to our definition $E_n^{\alpha_n}$ is the set consisting of $\phi(n)$ open intervals of length $2\alpha_n/n$ with centers at ν/n, where ν and n are relatively prime and $0 < \nu < n$. If (a, b) is some interval in $(0, 1)$, we shall need an estimate of the measure of the set common to $E_n^{\alpha_n}$ and the interval (a, b).

The number of intervals of $E_n^{\alpha_n}$ whose centers lie in $a < x \leqq b$ is exactly $\phi_b(n) - \phi_a(n)$. Then there are at least $\phi_b(n) - \phi_a(n) - 2$ intervals of $E_n^{\alpha_n}$ which lie entirely in (a, b) and at most $\phi_b(n) - \phi_a(n) + 2$ which can touch (a, b). Thus the measure of the set common to $E_n^{\alpha_n}$ and (a, b) is

$$(\phi_b(n) - \phi_a(n) + \theta)\frac{2\alpha_n}{n}$$

where $|\theta| \leqq 2$. Lemma III shows that this is equal to

$$\phi(n)(b - a + o(1))2\alpha_n/n = |E_n^{\alpha_n}|(b - a)(1 + o(1))$$

as $n \to \infty$, since $\phi(n) > \frac{1}{2}n^{\frac{1}{2}}$. The last $o(1)$ is less than $c_1 n^{-\frac{1}{4}}/(b - a)$, where c_1 is an absolute constant. From this follows

[6] A somewhat similar result is in Pólya-Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. I, p. 75, problem 188.

LEMMA IV. *If A is a given set in* (0, 1) *consisting of a finite number of intervals and* $\{\alpha_n\}$ *is a given sequence,* $0 \leqq \alpha_n < \frac{1}{2}$, *then*

$$| A \cdot E_n^{\alpha_n} | \leqq | A | | E_n^{\alpha_n} | (1 + cn^{-\frac{1}{2}}).$$

Here c is a constant which depends only on the set A.

Proof. We have proved Lemma IV in case A consists of a single interval, but in the general case A is the sum of a finite number of non-overlapping intervals

$$A = A_1 + A_2 + \cdots + A_k.$$

Then

$$| A_\nu E_n^{\alpha_n} | \leqq | A_\nu | | E_n^{\alpha_n} | \left(1 + \frac{c_1 n^{-\frac{1}{2}}}{| A_\nu |}\right);$$

so if $c = \max c_1/| A_\nu |$, we have

$$| AE_n^{\alpha_n} | \leqq \sum_{\nu=1}^{k} | A_\nu E_n^{\alpha_n} | \leqq | A | | E_n^{\alpha_n} | (1 + cn^{-\frac{1}{2}}).$$

The result follows.

Proof of Theorem I. Let the sequence $\{\alpha_n\}$ satisfy the conditions of Theorem I. For simplicity of notation let the sets $E_q^{\alpha_q}$ (which have been defined) be denoted by E_q. E_q consists of $\phi(q)$ open intervals, each of length $2\alpha_q/q$, with centers at p/q, where p and q are relatively prime and $0 < p < q$ if $q > 1$. Let

$$E = \sum_{2}^{\infty} E_q.$$

If it is shown that the measure of E is equal to 1, it will follow that for almost all x there is at least one set of relatively prime p and q such that

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q}. \tag{7}$$

We suppose that $| E | < 1$ and show that this leads to a contradiction.

Given a small positive number δ, let

$$A = E_2 + E_3 + \cdots + E_{q_1}$$

and choose q_1 so large that

$$| A | > | E | - \delta.$$

Then A consists of a finite number of intervals; so if q_2 is sufficiently large, we have from Lemma IV

$$| AE_q | \leqq | A | | E_q | (1 + \delta), \qquad q \geqq q_2(A, \delta). \tag{8}$$

Let n be greater than $q_1 + q_2$ and let

$$B = E_n + E_{n+1} + \cdots + E_m, \qquad m \geqq n > q_1 + q_2. \tag{9}$$

Now

$$\sum_{n}^{m} | E_\nu | \geqq | B | \geqq \sum_{n}^{m} | E_\nu | - \sum_{\nu=n+1}^{m} \sum_{j=n}^{\nu-1} | E_\nu E_j |;$$

so, by Lemma II,

$$| B | \geqq \sum_{\nu=n}^{m} | E_\nu | - 2\Big(\sum_{\nu=n}^{m} \alpha_\nu\Big)^2. \tag{10}$$

It follows from (8) that

$$| AB | \leqq \sum_{n}^{m} | AE_q | \leqq | A | \Big(\sum_{n}^{m} | E_q |\Big) (1 + \delta).$$

We have

$$| E | \geqq | A + B | = | A | + | B | - | AB |; \tag{11}$$

so, using (10) and (11), we have

$$| E | \geqq | A | + \Big\{\sum_{n}^{m} | E_\nu |\Big\}\{1 - | A | (1 + \delta)\} - 2\Big(\sum_{n}^{m} \alpha_\nu\Big)^2. \tag{12}$$

Choose δ so small that $| A |(1 + \delta) < 1$.

From the hypotheses of Theorem I it follows that there are arbitrarily large n and m such that

$$\sum_{\nu=n}^{m} \alpha_\nu > 1, \qquad \sum_{\nu=n}^{m} \frac{\alpha_\nu \phi(\nu)}{\nu} > \tfrac{1}{2} c \sum_{\nu=n}^{m} \alpha_\nu. \tag{13}$$

Then substituting in (12) we obtain

$$| E | \geqq | A | + c\{1 - | A | (1 + \delta)\} \Big\{\sum_{n}^{m} \alpha_\nu\Big\} - 2 \Big\{\sum_{n}^{m} \alpha_\nu\Big\}^2. \tag{14}$$

The right side is an expression of the form $| A | + bt - 2t^2$, where t represents $\sum \alpha_\nu$ and $b = c\{1 - | A |(1 + \delta)\}$ $(0 < b < 1)$. The maximum of this expression occurs when $t = \frac{1}{4}b$.

In order to satisfy the last condition ($t = \frac{1}{4}b$) we introduce the following device. Clearly if the length of some of the intervals is decreased, the measure of E will not be increased. Let z be a real number satisfying $0 < z < 1$. Let E'_q be the set in (0, 1) consisting of $\phi(q)$ open intervals, each of length $2z\alpha_q/q$, with centers at p/q, where p and q are relatively prime and $0 < p < q$. Clearly $E'_q \subset E_q$. Keeping A the same as before, we use in place of B the set

$$B_z = E'_n + E'_{n+1} + \cdots + E'_m.$$

Then (14) becomes

$$|E| \geqq |A| + c\{1 - |A|(1+\delta)\}\{\sum_n^m z\alpha_\nu\} - 2\{\sum_n^m z\alpha_\nu\}^2,$$

where $0 < z < 1$. Choose z such that $\sum_n^m z\alpha_\nu = \frac{1}{4}b$. Then we obtain

$$|E| \geqq |A| + \{\tfrac{1}{8}c^2\}\{1 - |A|(1+\delta)\}^2.$$

Letting $\delta \to 0$ we must have $|A| \to |E|$, so

$$|E| \geqq |E| + \{\tfrac{1}{8}c^2\}\{1 - |E|\}^2.$$

Hence $|E| = 1$.

It has thus been shown that for almost all x in $(0, 1)$, (7) is satisfied for at least one pair of relatively prime p and q.

To show that (7) is satisfied for arbitrarily many relatively prime p and q let m be some positive integer, and let $\{\alpha_q'\}$ be a new sequence defined by

$$\alpha_q' = \begin{cases} 0, & q \leqq m, \\ \alpha_q, & q > m. \end{cases}$$

Then the sequence $\{\alpha_q'\}$ satisfies all the conditions which in Theorem I are imposed on $\{\alpha_q\}$. Thus, by what we have just shown, for almost all x there is at least one set of relatively prime p and q such that

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q'}{q},$$

or what is the same, (7) is true for some $q > m$. Let D_m $(m = 1, 2, 3, \cdots)$ be the sets in $(0, 1)$ for which (7) is true for at least one pair of relatively prime p and q with $q > m$. Let D be the set common to all D_m. Then the measure of each D_m is equal to 1, so the measure of D is 1. If x lies in D, then (7) is true for arbitrarily many relatively prime p and q. This completes the proof of Theorem I.

From Theorem I we obtain

THEOREM II. *Let Q be an increasing sequence of positive integers such that $\phi(q)/q$ has a positive lower bound when $q \in Q$, and let $\{\alpha_q\}$ be a sequence of non-negative numbers.*

If $\sum_{q \in Q} \alpha_q = \infty$, then for almost all x there are arbitrarily many relatively prime p and q such that

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q}, \qquad q \in Q. \tag{15}$$

If $\sum_{q \in Q} \alpha_q < \infty$, then for almost no x are there arbitrarily many relatively prime p and q satisfying (15).

Proof. In view[7] of Theorem I it is only necessary to prove the last assertion. This will be a consequence of the following statement: If Q is an increasing sequence of positive integers and $\sum_{q \epsilon Q} \alpha_q \phi(q)/q$ converges, $\alpha_q \geqq 0$, then the set of points for which there are arbitrarily many relatively prime p and q such that

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q}, \qquad q \epsilon Q, \tag{16}$$

is of measure zero.

To show this let A be the set of points in (0, 1) for which (16) is satisfied for arbitrarily many relatively prime p and q. Then if x lies in A, it must lie in arbitrarily many $E_q^{\alpha_q}$; hence for every m

$$A \subset \sum_{q > m} E_q^{\alpha_q}.$$

But $| E_q^{\alpha_q} | = 2\alpha_q \phi(q)/q$, so

$$| A | \leqq \sum_{q > m} 2\alpha_q \phi(q)/q,$$

and this tends to zero as m tends to infinity.

If $m_1, m_2, m_3, \cdots$ is an increasing sequence of positive integers, and if

$$\lim_{m \to \infty} \frac{1}{m} \sum_{m_\nu \leqq m} 1$$

exists, then the sequence $\{m_\nu\}$ is said to have a density (frequency), which is then equal to the limit. If the limit does not exist, then the upper and lower limits of indetermination of the same expression are called the upper and lower densities respectively.

According to Schoenberg[8] for "most" integers the function $\phi(n)/n$ is appreciably greater than zero. More precisely, for every γ $(0 < \gamma < 1)$ there is a $\delta > 0$ such that the density of the numbers for which

$$\frac{\phi(n)}{n} > \delta$$

is greater than γ.

Suppose that Q is an increasing sequence of positive integers and $\{\alpha_q\}$ is a sequence of non-negative numbers. We make no explicit supposition regarding $\phi(q)/q$. The problem is to find conditions under which for almost all x

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q}, \qquad q \epsilon Q, \tag{17}$$

[7] Theorems I and II are essentially equivalent. For Theorem II follows from Theorem I. Conversely, if $\{\alpha_q\}$ satisfies conditions (a) and (b″) of Theorem I, there will be a sequence of integers Q such that $\sum_{q \epsilon Q} \alpha_q \phi(q)/q$ diverges and $\phi(q)/q$ has a positive lower bound for $q \epsilon Q$. This sequence will satisfy the conditions of Theorem II.

[8] I. J. Schoenberg, *On asymptotic distributions of arithmetical functions*, Transactions of the American Mathematical Society, vol. 39(1936), pp. 315–330.

for arbitrarily many relatively prime p and q. If the sequence Q has a positive lower density, then there will be a subsequence A of Q for which $\phi(q)/q$ has a positive lower bound. If the α_q are not too irregular (for example, if α_q is decreasing), then we show that the subsequence A may be so chosen that $\sum_{q \in A} \alpha_q$ diverges. It will follow from Theorem II that (17) is satisfied (it will in fact be satisfied for $q \in A$). More generally, we have

THEOREM III. *Let Q be an increasing sequence of integers with a positive lower density. Let $\alpha_1, \alpha_2, \alpha_3, \cdots$ be a sequence of positive numbers such that*

$$\sum_{q=1}^{\infty} \alpha_q = \infty$$

and, for some real c, α_q/q^c is a decreasing function of q. Then for almost all x there exist arbitrarily many relatively prime p and q such that

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q}, \qquad q \in Q.$$

Proof. Let the lower density of the sequence $Q = q_1, q_2, \cdots$ be λ $(0 < \lambda \leqq 1)$. If ρ is sufficiently small, the density of the integers for which

$$\frac{\phi(n)}{n} > \rho$$

is greater than $1 - \frac{1}{2}\lambda$. Let A be the subsequence of Q for which $\phi(q)/q > \rho$. Then A is a sequence common to one of density greater than $1 - \frac{1}{2}\lambda$ and one of lower density λ. The lower density of A is therefore greater than $\frac{1}{2}\lambda$.

Let

$$\delta_q = \begin{cases} 1 & \text{if } q \in A, \\ 0 & \text{otherwise.} \end{cases}$$

If α_q/q^c decreases as q increases, then so does α_q/q^k for any $k > c$. We therefore suppose without loss of generality that c is an integer greater than zero. Let

$$\sigma_m = \sum_{q=1}^{m} \delta_q q^c, \qquad \sigma_0 = 0.$$

Then

$$\sigma_m \geqq \sum_{q \geqq \lambda m/4}^{m} \delta_q q^c \geqq (\tfrac{1}{4}\lambda m)^c \sum_{q \geqq \lambda m/4}^{m} \delta_q$$
$$\geqq (\tfrac{1}{4}\lambda m)^c \left\{ \sum_{1}^{m} \delta_q - \tfrac{1}{4}\lambda m \right\}.$$

The last sum is the number of integers of A which are equal to or less than m; and so this sum is greater than $\frac{1}{2}\lambda m$ for all large m. If m_1 is large enough,

$$\sigma_m > (\tfrac{1}{4}\lambda m)^{c+1}, \qquad m \geqq m_1 .$$

Using Abel's transformation, we obtain, when $m > m_1$,

$$\begin{aligned}\sum_1^m \alpha_q \delta_q &= \sum_1^m \alpha_q(\sigma_q - \sigma_{q-1})/q^c \\ &= \sum_1^{m-1} \sigma_q\{\alpha_q/q^c - \alpha_{q+1}/(q+1)^c\} + \sigma_m \alpha_m/m^c \\ &\geqq (\tfrac{1}{4}\lambda)^{c+1}\left\{\sum_{m_1}^{m-1} q^{c+1}(\alpha_q/q^c - \alpha_{q+1}/(q+1)^c) + m\alpha_m\right\} \\ &\geqq (\tfrac{1}{4}\lambda)^{c+1}\sum_{m_1}^{m} \alpha_q(q^{c+1} - (q-1)^{c+1})/q^c \\ &\geqq (\tfrac{1}{8}\lambda)^{c+1}\sum_{m_1}^{m} \alpha_q\end{aligned}$$

since by the mean value theorem $q^{c+1} - (q-1)^{c+1} > (c+1)(q-1)^c$, which is greater than $(\frac{1}{2}q)^c$ for $q > 2$.

It now follows that $\sum \alpha_q \delta_q$ is a diverging series of non-negative terms, or what is the same, $\sum_{q \epsilon A} \alpha_q = \infty$. But $\phi(q)/q > \rho$ if q belongs to A, so Theorem II is immediately applicable.

To prove the theorem of Walfisz mentioned earlier suppose that

(a) $\sum_{q=1}^{\infty} \alpha_q = \infty, \qquad \alpha_q > 0;$

(b) *$q\alpha_q$ is a decreasing function of q.*

Let Q be the sequence 4, 8, 12, 16, $\cdots$ of positive integers. Since the density of Q is $\frac{1}{4}$, Theorem III is immediately applicable. It states that for almost all x there exist arbitrarily many relatively prime integers p and q such that

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q},$$

where $q \equiv 0 \pmod 4$.

We now turn to the construction of an example which shows that the mere divergence of $\sum \alpha_\nu$ (condition (a)) is not sufficient to insure the conclusions of Khintchine's theorem. The example depends on the following lemma.

Lemma V. *Let R and ϵ be given positive numbers (R large and ϵ small). There is a finite sequence $\{\alpha_q\}$ of non-negative numbers such that*

$$\sum \alpha_q > 1; \qquad \alpha_q = 0, \text{ when } q \leqq R,$$

but for x in $(0, 1)$ the inequality

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q} \tag{18}$$

can be satisfied only in a set of measure less than ϵ.

Proof. Let α be some positive number less than $\frac{1}{2}\epsilon$ and let

$$N = p_1 p_2 \cdots p_k,$$

where p_i are different primes, each greater than R, and let

$$\prod_{i=1}^{k}\left(1+\frac{1}{p_i}\right) > 1 + \frac{1}{2\alpha}.$$

This is possible since the product on the left when extended over all primes diverges. Let

$$\alpha_q = \begin{cases} q\alpha/N & \text{if } q > 1 \text{ and is a divisor of } N, \\ 0 & \text{otherwise,} \end{cases}$$

and let E_q be the set in $(0, 1)$ consisting of $q - 1$ open intervals of length $2\alpha_q/q$ with centers at p/q ($p = 1, 2, 3, \cdots, q - 1$), and the two open intervals $(0, \alpha_q/q)$ and $(1 - \alpha_q/q, 1)$. Then $|E_q| = 2\alpha_q = 2q\alpha/N$. If

$$E = \sum E_q,$$

it is clear that the sets E_q are all included in the set E_N, so we have $E = E_N$. Inequality (18) will be satisfied for some x ($0 < x < 1$) if and only if x belongs to the set E. Thus it need only be shown that $\sum |E_q| > 1$ and that $|E| = |E_N| < \epsilon$. The second inequality follows from $\alpha < \frac{1}{2}\epsilon$. Using a well-known expression for the sum of the divisors of an integer, we obtain

$$\sum |E_q| = \frac{2\alpha}{N}\sum_{q|N} q = \frac{2\alpha}{N}\left\{\prod(1 + p_i) - 1\right\},$$

and this is greater than

$$2\alpha\left\{\prod\left(1+\frac{1}{p_i}\right) - 1\right\} > 1.$$

The required example is now easily constructed. The statement is:

There is a sequence $\{\alpha_q\}$ of non-negative numbers such that $\sum_{q=1}^{\infty} \alpha_q = \infty$, but the set of points for which

$$\left| x - \frac{p}{q}\right| < \frac{\alpha_q}{q} \tag{19}$$

is satisfied for arbitrarily many integers p and q is of measure zero.

The p and q are not required to be relatively prime in this example.

Let the finite sequence $\{\alpha_q^{(1)}\}$ satisfy Lemma V with $R_1 = 1$, $\epsilon = 2^{-1}$. Then for some R_2, $\alpha_q^{(1)} = 0$ when $q \geqq R_2$. Let $\{\alpha_q^{(2)}\}$ satisfy Lemma V with $R = R_2$, $\epsilon = 2^{-2}$. Clearly this process may be continued indefinitely. We obtain a

sequence $\{\alpha_q\} = \{\alpha_q^{(1)}\} + \{\alpha_q^{(2)}\} + \{\alpha_q^{(3)}\} + \cdots$ of non-negative numbers satisfying

$$\sum_{q=1}^{\infty} \alpha_q = \infty.$$

But for x in $(0, 1)$ the inequality

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q}, \qquad q > R_k,$$

can be satisfied only if x belongs to a set of measure at most $\sum_{\nu=k}^{\infty} 2^{-\nu} = 2^{-k+1}$. The assertion follows.

We have left unsolved the problem of finding necessary and sufficient conditions on the sequence $\{\alpha_q\}$, $\alpha_q \geqq 0$, in order that for almost all x there exist arbitrarily many p and q satisfying

$$\left| x - \frac{p}{q} \right| < \frac{\alpha_q}{q}.$$

A first guess might be that a necessary and sufficient condition is the divergence of $\sum \alpha_q$, but the last example shows that this is not so. Turning to approximation by reduced fractions (p and q relatively prime), it is tempting to suppose that a necessary and sufficient condition is the divergence of $\sum \alpha_q \phi(q)/q$. We have been unable to decide upon this question.

UNIVERSITY OF ILLINOIS AND STANFORD UNIVERSITY.

NON-ALTERNATING AND NON-SEPARATING TRANSFORMATIONS MODULO A FAMILY OF SETS

By John W. Odle

1. **Introduction.** In a previous paper, by E. P. Vance [5],[1] weakly non-separating, weakly non-alternating, weakly completely non-separating, and weakly completely non-alternating transformations have been defined and discussed. It is the purpose of this paper to study some natural generalizations of these four transformations and to obtain Vance's types as one of the special cases. In the above transformations special favor was allotted to sets consisting of but a single point by allowing them to separate, whereas in the generalized transformations which are to be defined here, finite sets, continua, and other types of sets will be permitted to effect separations. This is the basis of the generalization accomplished in this paper. It will be seen that many of the same theorems which Vance obtained will go through in the more general situation also.

In the last two sections of this paper locally non-separating and locally non-alternating transformations, which were also defined by Vance in his paper, will be studied further, and some relationships between 0-regular transformations and locally non-alternating transformations will be established.

In this paper all transformations, $T(A) = B$, are assumed to be single valued and continuous, and the spaces considered are assumed to be compact metric continua.

The four new types of continuous transformations which are to be studied in this paper are defined as follows:

(1) T is called *non-separating modulo* $\mathcal{G}$ if, for any point x of B, the set $T^{-1}(x)$ does not separate two points in A unless a subset of $T^{-1}(x)$ which belongs to the collection $\mathcal{G}$ also separates the same two points in A.

(2) T is called *non-alternating modulo* $\mathcal{G}$ if, for any two points x and y of B, the set $T^{-1}(x)$ does not separate two points of $T^{-1}(y)$ in A unless a subset of $T^{-1}(x)$ which belongs to the collection $\mathcal{G}$ also separates the same two points in A.

(3) T is called *completely non-separating modulo* $\mathcal{G}$ if, for any point x of B and any closed subset M of $T^{-1}(x)$, M does not separate two points in A unless a subset of M which belongs to the collection $\mathcal{G}$ also separates the same two points in A.

Received August 20, 1940; presented to the American Mathematical Society, November 23, 1940. The writer wishes to express his appreciation for the help given by Professor W. L. Ayres of the University of Michigan, who suggested the problem and gave valuable criticism and advice during the preparation of this paper.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

(4) T is called *completely non-alternating modulo* $\mathcal{S}$ if, for any two points x and y of B and any closed subset M of $T^{-1}(x)$, M does not separate two points of $T^{-1}(y)$ in A unless a subset of M which belongs to the collection $\mathcal{S}$ also separates the same two points in A. (x and y may be identical here.)

If the null set is taken to be the only set in the collection $\mathcal{S}$, then the preceding definitions reduce to the definitions of non-separating, non-alternating, completely non-separating, and completely non-alternating transformations, which have been studied in considerable detail in previous papers by other authors [7, 9].

If the collection $\mathcal{S}$ is defined to be the totality of all single-point sets of the space A, then the definitions given here reduce to the definitions of weakly non-separating, weakly non-alternating, weakly completely non-separating, and weakly completely non-alternating transformations, which were given by Vance [5] in his recent paper.

Other collections $\mathcal{S}$ which will be mentioned later are the class of continua, the class of finite sets, and the class of sets having a finite number of components. Other meanings are of course possible, and perhaps useful. The beauty of the definitions of the transformations given here lies in their generality and applicability to many special cases.

It should be noted that if $\mathcal{S}$ is the collection of all closed sets, then any continuous transformation satisfies all four of the definitions which we have given above. This follows directly from the definitions and from the fact that $T^{-1}(x)$ is a closed set, for any x. This means, of course, that such a liberal definition for the collection $\mathcal{S}$ is unproductive of interesting results.

Certain relations exist between the transformations here defined, independently of the definition given to the collection $\mathcal{S}$. These relations are: If T is completely non-separating modulo $\mathcal{S}$, then T is completely non-alternating modulo $\mathcal{S}$; if T is completely non-separating modulo $\mathcal{S}$, then T is non-separating modulo $\mathcal{S}$; if T is completely non-alternating modulo $\mathcal{S}$, then T is non-alternating modulo $\mathcal{S}$; if T is non-separating modulo $\mathcal{S}$, then T is non-alternating modulo $\mathcal{S}$. These relations follow directly from the definitions given for the various types of transformations. By means of simple examples easily constructible for any of the specific collections $\mathcal{S}$ which we have considered, it can be seen that these are the only relations of implication which always subsist between these transformations, independently of the definition used for collections $\mathcal{S}$.

2. **Some characteristic properties.** Vance proved a number of theorems which delineated the properties of the transformations which he defined. Most of these theorems are still valid when stated for the general modulo $\mathcal{S}$ transformations with the collection $\mathcal{S}$ unspecified. The statements of these theorems will now be given, but the proofs will be omitted, as they follow the pattern of Vance's proofs with very slight modification. These theorems hold independently of the definition given to collection $\mathcal{S}$.

THEOREM 2.1. *If the transformation T is non-separating modulo $\mathcal{S}$, then, for any cut point b of B, the set $T^{-1}(b)$ contains a set K of the collection $\mathcal{S}$ which separates the space A.*

THEOREM 2.2. *If T is completely non-alternating modulo $\mathcal{S}$, then every open set contained in any set $T^{-1}(x)$ must contain a set K of the collection $\mathcal{S}$ which is a cut set of A.*

COROLLARY 2.21. *If T is completely non-separating modulo $\mathcal{S}$, then every open set contained in any set $T^{-1}(x)$ must contain a cut set of A which is in the collection $\mathcal{S}$.*

THEOREM 2.3. *If T is completely non-separating modulo $\mathcal{S}$, and if b is a cut point of B, and if $T^{-1}(b)$ contains an irreducible cut set M between two points of A, then M is a set of the collection $\mathcal{S}$.*

THEOREM 2.4. *Let A, and therefore B, be locally connected. If T is non-alternating and if, for any cut point b of B, every subset of $T^{-1}(b)$ which is an irreducible cut set of A between two points of A is a set of the collection $\mathcal{S}$, then T is non-separating modulo $\mathcal{S}$.*

THEOREM 2.5. *Let A, and therefore B, be locally connected. If T is completely non-alternating and if, for any cut point b of B, every subset of $T^{-1}(b)$ which is an irreducible cut set of A between two points of A is a set of the collection $\mathcal{S}$, then T is completely non-separating modulo $\mathcal{S}$.*

3. **Product and factor theorems.** With the purpose of obtaining a complete analysis of all product and factor theorems involving the following five types of transformations: monotone, non-separating, non-alternating, non-separating modulo $\mathcal{S}$ and non-alternating modulo $\mathcal{S}$, where the collection $\mathcal{S}$ is the totality of all single-point sets, we have set down all possible combinations of these transformations and investigated every one to see if it would lead to a true theorem. As a result of this systematic investigation, it can now be stated definitely that the previously known product and factor theorems, given by Vance [5], Wardwell [7], and Whyburn [9], plus two new ones included in this paper for the first time, are the only true product and factor theorems possible involving the above-mentioned five types of transformations. Every other combination has been shown by an example to lead to a false theorem.

Let $T_1(A) = B$ and $T_2(B) = C$. Then by $T = T_2 \cdot T_1$ we mean $T(A) = T_2[T_1(A)] = T_2(B) = C$.

The two new theorems are as follows:

THEOREM 3.1. *Let the collection $\mathcal{S}$ be the totality of all single-point sets. If T_1 is monotone and non-separating modulo $\mathcal{S}$ and if T_2 is non-separating modulo $\mathcal{S}$, then T is also non-separating modulo $\mathcal{S}$.*

Proof. Suppose T is not non-separating modulo $\mathcal{S}$; i.e., assume there is a point x in C such that $T^{-1}(x)$ separates two points a_1 and a_2 in A and no single point of $T^{-1}(x)$ does this. Let $A - T^{-1}(x) = A_1 + A_2$ be a separation, where

A_1 contains a_1 and A_2 contains a_2. Consider $B_1 = T_1(A_1)$ and $B_2 = T_1(A_2)$. Since T_1 is monotone, $B_1 \cdot B_2 = 0$. Then it follows from the continuity of T_1 and the compactness of A that B_1 and B_2 are separated sets. Hence $B - T_2^{-1}(x) = B_1 + B_2$ is a separation of B, where B_1 contains $T_1(a_1)$ and B_2 contains $T_1(a_2)$. Since T_2 is non-separating modulo $\mathcal{G}$, some single point r of $T_2^{-1}(x)$ must separate $T_1(a_1)$ and $T_1(a_2)$ in B. Thus $B - r = B_3 + B_4$ is a separation, where B_3 contains $T_1(a_1)$ and B_4 contains $T_1(a_2)$. Then $A - T_1^{-1}(r) = T_1^{-1}(B_3) + T_1^{-1}(B_4)$ is a separation of A in which the first set contains a_1 and the second contains a_2. Since T_1 is non-separating modulo $\mathcal{G}$, some single point y of $T_1^{-1}(r)$ must also separate a_1 and a_2 in A. But then y is contained in $T^{-1}(x)$, and this contradicts our first assumption. Hence T must be non-separating modulo $\mathcal{G}$.

THEOREM 3.2. *Let the collection $\mathcal{G}$ be the totality of all single-point sets. If T_1 is monotone and non-separating modulo $\mathcal{G}$, and if T_2 is non-alternating modulo $\mathcal{G}$, then T is also non-alternating modulo $\mathcal{G}$.*

Proof. Suppose T is not non-alternating modulo $\mathcal{G}$; i.e., assume there are two points x and y in C such that $T^{-1}(x)$ separates two points a_1 and a_2 of $T^{-1}(y)$ in A, and no single point of $T^{-1}(x)$ does this. Let $A - T^{-1}(x) = A_1 + A_2$ be a separation of A, where A_1 contains a_1 and A_2 contains a_2. Consider $B_1 = T_1(A_1)$ and $B_2 = T_1(A_2)$. As in the preceding theorem, B_1 and B_2 are separated sets, and $B - T_2^{-1}(x) = B_1 + B_2$, where B_1 contains $T_1(a_1)$ and B_2 contains $T_1(a_2)$. The sets $T_1(a_1)$ and $T_1(a_2)$ are contained in $T_2^{-1}(y)$, and since T_2 is non-alternating modulo $\mathcal{G}$, it follows that $T_2^{-1}(x)$ contains some single point r which also separates $T_1(a_1)$ and $T_1(a_2)$ in B. Thus $B - r = B_3 + B_4$ is a separation, where B_3 contains $T_1(a_1)$ and B_4 contains $T_1(a_2)$. Then $A - T_1^{-1}(r) = T_1^{-1}(B_3) + T_1^{-1}(B_4)$ is a separation of A in which the first set contains a_1 and the second contains a_2. Since T_1 is non-separating modulo $\mathcal{G}$, some single point s of $T_1^{-1}(r)$ must separate a_1 and a_2 in A. But then s is contained in $T^{-1}(x)$, and this contradicts the first assumption. Hence T must be non-alternating modulo $\mathcal{G}$.

The problem of generalizing these product and factor theorems so that they would hold for more general collections $\mathcal{G}$ was considered, and we found that in most cases such severe restrictions had to be imposed that the theorems obtained were of no importance or interest. However, two theorems of Vance's are generalizable with only slight restrictions, and they give the following theorems which hold for collections $\mathcal{G}$ unrestricted except for the condition stated in the hypothesis.

THEOREM 3.3. *If the collection $\mathcal{G}$ is closed under continuous transformations*[2] *and if T is non-separating modulo $\mathcal{G}$ and T_1 is monotone, then T_2 is non-separating modulo $\mathcal{G}$.*[3]

[2] DEFINITION. The collection $\mathcal{G}$ is said to be *closed under continuous transformations* if the continuous transform of any set in the collection is also a set of the collection.

[3] This theorem may be stated in a somewhat more general fashion as follows: If T is non-separating modulo $\mathcal{G}$ and T_1 is monotone, then T_2 is non-separating modulo $T_1(\mathcal{G})$, where $T_1(\mathcal{G})$ is the collection of sets in B which are images under T_1 of those sets in A which belong to $\mathcal{G}$. The same proof is valid for either statement of the theorem.

Proof. Assume T_2 is not non-separating modulo $\mathcal{S}$; i.e., assume there is a point x of C such that $T_2^{-1}(x)$ separates two points b_1 and b_2 of B and $T_2^{-1}(x)$ contains no set of the collection $\mathcal{S}$ which does this. Let $B - T_2^{-1}(x) = B_1 + B_2$ be a separation of B, where B_1 contains b_1 and B_2 contains b_2. Then $A - T^{-1}(x) = T_1^{-1}(B_1) + T_1^{-1}(B_2)$ is a separation of A, in which $T_1^{-1}(B_1)$ contains $T_1^{-1}(b_1)$, and $T_1^{-1}(B_2)$ contains $T_1^{-1}(b_2)$. Let a_1 be a point of $T_1^{-1}(b_1)$ and a_2 be a point of $T_1^{-1}(b_2)$. Then, since T is non-separating modulo $\mathcal{S}$, the set $T^{-1}(x)$ contains a set K of the collection $\mathcal{S}$ which separates a_1 and a_2 in A. Let $A - K = A_1 + A_2$ be a separation, where A_1 contains a_1 and A_2 contains a_2. Consider the sets $[T_1(A_1) - T_1(K)]$ and $[T_1(A_2) - T_1(K)]$. The first set in brackets contains the point b_1 because the point a_1 of $T_1^{-1}(b_1)$ is contained in A_1, and $T_1(K)$ is a subset of $T_2^{-1}(x)$, which does not contain b_1. Likewise, the second set in brackets contains the point b_2. Since T_1 is monotone, these two sets are disjoint; and then because of the compactness of A and the continuity of T_1, it follows that they are separated sets. Thus we have $B - T_1(K) = [T_1(A_1) - T_1(K)] + [T_1(A_2) - T_1(K)]$ is a separation of B. The set K is a set of the collection $\mathcal{S}$ and therefore, by hypothesis, $T_1(K)$ is also a set of the collection $\mathcal{S}$. But then the separation of B by $T_1(K)$ contradicts our first assumption. Hence T_2 must be non-separating modulo $\mathcal{S}$.

THEOREM 3.4. *If the collection $\mathcal{S}$ is closed under continuous transformations, and if T is non-alternating modulo $\mathcal{S}$ and T_1 is monotone, then T_2 is non-alternating modulo $\mathcal{S}$.*[4]

Proof. Assume T_2 is not non-alternating modulo $\mathcal{S}$; i.e., assume there are points x and y of C such that $T_2^{-1}(x)$ separates two points b_1 and b_2 of $T_2^{-1}(y)$ in B, and $T_2^{-1}(x)$ contains no set of the collection $\mathcal{S}$ which does this. Let $B - T_2^{-1}(x) = B_1 + B_2$ be a separation, where B_1 contains b_1 and B_2 contains b_2. Then $A - T^{-1}(x) = T_1^{-1}(B_1) + T_1^{-1}(B_2)$ is a separation of A in which $T_1^{-1}(B_1)$ contains $T_1^{-1}(b_1)$ and $T_1^{-1}(B_2)$ contains $T_1^{-1}(b_2)$. Let a_1 be a point of $T_1^{-1}(b_1)$ and a_2 be a point of $T_1^{-1}(b_2)$. Then a_1 and a_2 are points of $T^{-1}(y)$, and since T is non-alternating modulo $\mathcal{S}$, the set $T^{-1}(x)$ contains a set K of the collection $\mathcal{S}$ which separates a_1 and a_2 in A. Let $A - K = A_1 + A_2$ be a separation of A, where A_1 contains a_1 and A_2 contains a_2. Then, just as in the preceding theorem, we have $B - T_1(K) = [T_1(A_1) - T_1(K)] + [T_1(A_2) - T_1(K)]$ is a separation of B. The first set in brackets contains the point b_1 because the point a_1 of $T_1^{-1}(b_1)$ is contained in A_1, and $T_1(K)$ is a subset of $T_2^{-1}(x)$, which does not contain b_1. Likewise, the second set in brackets contains the point b_2. The set K is a set of the collection $\mathcal{S}$ and therefore, by hypothesis, $T_1(K)$ is also a set of the collection $\mathcal{S}$. But then the separation of B by $T_1(K)$ contradicts our first assumption. Hence T_2 must be non-alternating modulo $\mathcal{S}$.

The condition placed in the hypotheses of the two preceding theorems, that

[4] This theorem may also be stated in a more general fashion: *If T is non-alternating modulo $\mathcal{S}$ and T_1 is monotone, then T_2 is non-alternating modulo $T_1(\mathcal{S})$.*

the continuous transformation of a set of the collection $\mathcal{G}$ is again a set of the collection $\mathcal{G}$, seems to be a rather weak one, and hence it does not restrict seriously the applications of the theorem. The condition is satisfied by all the collections $\mathcal{G}$ which we have had in mind; i.e., collections consisting of single points, finite point sets, connected sets, and sets having a finite number of components.

4. **Applications involving special collections $\mathcal{G}$.** In this section various particular collections $\mathcal{G}$ will be considered, and the behavior of the transformations modulo $\mathcal{G}$ for these definitions of $\mathcal{G}$ will be studied on certain special curves and surfaces.

THEOREM 4.1. *If A is an n-coherent Peano continuum,*[5] *then any continuous transformation T on A is completely non-separating modulo $\mathcal{G}$, where the collection $\mathcal{G}$ is the family of all sets consisting of n or fewer components.*

Proof. Suppose the continuous transformation T is not completely non-separating modulo $\mathcal{G}$; i.e., assume there is a point x of B and a closed set M contained in $T^{-1}(x)$ such that M separates two points a_1 and a_2 in A and M contains no set of the collection $\mathcal{G}$ which does this. Let $A - M = A_1 + A_2$ be a separation, where A_1 contains a_1 and A_2 contains a_2. Denote by C the component of A_1 containing a_1. Then a_2 is contained in $A - \bar{C}$. Let D be the component of $A - \bar{C}$ containing a_2. The sets C and D are thus both open sets, since in Peano space components of open sets are open. The set $A - D$ is connected (see [1]). Then $\overline{A - D}$ and $\bar{D}$ are continua. Obviously $A = \bar{D} + \overline{A - D}$. Therefore, since A is n-coherent, $\bar{D}\cdot\overline{A - D}$ consists of n or fewer components, i.e., $\bar{D}\cdot\overline{A - D}$ is a set of the collection $\mathcal{G}$.

Let $F = \bar{D}\cdot\overline{A - D}$. Then F is the frontier of D. Further, F is contained in $\bar{D} - D$, because no point of the open set D can be a limit point of its complement $A - D$. Also $\bar{D} - D$ is contained in $\bar{C}$, since D is a component of $A - \bar{C}$, and any limit point of D which was not in $\bar{C}$ would have been added to D. Furthermore, since C is an open set and D is in $A - C$, the set C contains no limit points of D. Hence $\bar{D} - D$ is contained in $\bar{C} - C$. Also $\bar{C} - C$ is contained in M, as the following reasoning shows. Consider any limit point y of C. The point y is not an element of A_2, since C is in A_1 which is separated from A_2. If y is an element of A_1, then it would be contained in the component C. Thus the only other possibility is that y is an element of M. Bringing all these facts together we have:

$$F = \bar{D}\cdot\overline{A - D} \subset \bar{D} - D \subset \bar{C} - C \subset M.$$

Now $A - F = (\bar{D} + \overline{A - D}) - F = (\bar{D} - F) + (\overline{A - D} - F)$. These two sets in parentheses are separated, because when the common part of two

[5] DEFINITION. The continuum A is said to be *n-coherent* if, whenever A is expressed as the sum of two continua, $A = A_1 + A_2$, then $A_1 \cdot A_2$ consists of n or fewer components. In the case in which $n = 1$, A is called *unicoherent*.

continua is subtracted out, neither remainder can contain a limit point of the other. Now $\bar{D} - F$ contains a_2, since a_2 is in the interior of D, and F is contained in $\bar{D} - D$. Also, $\overline{A - D} - F$ contains a_1, because D is contained in $A - \bar{C}$, hence $\bar{C} \subset A - D \subset \overline{A - D}$; and a_1 is an element of C, and F is contained in $\bar{C} - C$.

Thus we have contained in M a subset F of the collection $\mathcal{G}$ which separates a_1 and a_2 in A. This is a contradiction. Hence T must be completely non-separating modulo $\mathcal{G}$.

This theorem is not true in non-Peanian spaces, as the following example shows for the case when $n = 1$. Let A consist of all points in the analytic plane satisfying one of the following conditions:

(a) $x = 0, 0 \leqq y \leqq 1$;

(b) $y = \frac{1}{i}, 0 \leqq x \leqq 1, i = 1, \cdots, n, \cdots$;

(c) $y = 0, 0 \leqq x \leqq 1$.

A is then a unicoherent continuum. Define the transformation T to be one which sends all points of A which have an abscissa equal to $\frac{1}{2}$ into the point $(\frac{1}{2}, 0)$ and let T be a homeomorphism elsewhere. This transformation is continuous. But the inverse of the point $(\frac{1}{2}, 0)$ separates the two points $(1, 0)$ and $(0, 0)$ in A, and no single component of the inverse effects this separation. Hence T is not non-separating modulo $\mathcal{G}$, where the collection $\mathcal{G}$ is the family of all sets consisting of one component, i.e., connected sets.

Theorem 4.2. *Let the collection $\mathcal{G}$ be any family of closed subsets of A which includes all single-point sets. A necessary and sufficient condition that no set of the collection $\mathcal{G}$ separate the space A is that every transformation on A which is non-separating modulo $\mathcal{G}$ be also non-separating in the strict sense.*

Proof. Necessity. Suppose A has the property that no set of the collection $\mathcal{G}$ separates A, but assume that there is a transformation T which is non-separating modulo $\mathcal{G}$ but which is not non-separating in the strict sense. This means that there is a point x of B such that $T^{-1}(x)$ separates two points a_1 and a_2 in A, and $T^{-1}(x)$ contains a set of the collection $\mathcal{G}$ which also separates a_1 and a_2 in A. But this contradicts the hypothesis that no set of the collection $\mathcal{G}$ separates A. Hence T must be non-separating in the strict sense.

Sufficiency. Suppose that any transformation on A which is non-separating modulo $\mathcal{G}$ is also non-separating in the strict sense. Now assume that some set M of the collection $\mathcal{G}$ separates A. Then, since M is a closed set, the decomposition of A which has M as one element and all other points in A individually as elements is upper semi-continuous [4], and hence this decomposition defines a continuous transformation T on A which sends M into one point x and is a homeomorphism elsewhere on A (cf. [2]). Now every inverse set $T^{-1}(y)$ belongs to the collection $\mathcal{G}$, for $T^{-1}(y)$ is either a single point of A or the set M and hence belongs to $\mathcal{G}$ in either case. Hence every inverse set $T^{-1}(y)$ which separates A contains a separating subset (namely, itself) belonging to $\mathcal{G}$, and thus T

is non-separating modulo $\mathcal{G}$. But T is not non-separating in the strict sense, for $T^{-1}(x) = M$ separates A. This contradicts our hypothesis and hence A must have the property that no set of the collection $\mathcal{G}$ separates A.

The following theorems are stated without proof, because the proofs are very similar to the one given for Theorem 4.2.

THEOREM 4.3. *Let $\mathcal{G}$ be any family of closed subsets of A which includes all subsets consisting of one or two points. In order that no set of the collection $\mathcal{G}$ separate A it is necessary and sufficient that every transformation on A which is non-alternating modulo $\mathcal{G}$ be also non-alternating in the strict sense.*

THEOREM 4.4. *Let $\mathcal{G}$ be any hereditary family*[6] *of closed subsets of A which includes all single-point subsets. In order that no set of the family $\mathcal{G}$ separate A it is necessary and sufficient that every transformation on A which is completely non-separating modulo $\mathcal{G}$ be also completely non-separating in the strict sense.*

THEOREM 4.5. *Let $\mathcal{G}$ be any hereditary family of closed subsets of A which includes all sets consisting of one or two points. In order that no set of the family $\mathcal{G}$ separate A it is necessary and sufficient that every transformation on A which is completely non-alternating modulo $\mathcal{G}$ be also completely non-alternating in the strict sense.*

A number of corollaries to the preceding theorems are easily deduced, and a few of them will be stated here. These also are given without proof, for either they are just special cases of the preceding theorems, or their proofs are very similar to the proof of Theorem 4.2.

COROLLARY. *A necessary and sufficient condition that a Peano continuum A be a simple closed curve is that any transformation on A which is* $\left\{\begin{array}{l}\textit{non-separating}\\ \textit{non-alternating}\end{array}\right\}$ *modulo $\mathcal{G}$, where $\mathcal{G}$ is the collection of all continua in A, be also* $\left\{\begin{array}{l}\textit{non-separating}\\ \textit{non-alternating}\end{array}\right\}$ *in the strict sense.*

This corollary makes use of the theorem that a Peano continuum is a simple closed curve if and only if it is not separated by any subcontinuum (cf. [3]).

COROLLARY. *Let the collection $\mathcal{G}$ consist of all single-point sets in A. In order that the space A have no cut point it is necessary and sufficient that every transformation T on A which is* $\left\{\begin{array}{l}\textit{completely non-separating}\\ \textit{completely non-alternating}\\ \textit{non-separating}\\ \textit{non-alternating}\end{array}\right\}$ *modulo $\mathcal{G}$ be also* $\left\{\begin{array}{l}\textit{completely non-separating}\\ \textit{completely non-alternating}\\ \textit{non-separating}\\ \textit{non-alternating}\end{array}\right\}$ *in the strict sense.*

[6] DEFINITION. A family $\mathcal{G}$ of closed sets is said to be hereditary if every closed subset of any set of the family $\mathcal{G}$ also belongs to $\mathcal{G}$.

5. **Locally non-separating transformations.** Locally non-separating transformations were defined by Vance [5] as follows:

A continuous transformation T is called *locally non-separating at a point* p of A provided for any neighborhood U_p of p there exists a neighborhood V_p of p, contained in U_p, such that if u and v are two points of $V_p - T^{-1}[T(p)]$, then u and v lie in a connected subset of $U_p - T^{-1}[T(p)]$. A continuous transformation is called *locally non-separating on the space* A if it is locally non-separating at every point of A.

Vance gave examples to show that non-separating transformations need not be locally non-separating, and that locally non-separating transformations need not be non-separating.

THEOREM 5.1. *If A contains an open set U such that* (a) *$\bar{U}$ is a Peano continuum,* (b) *every non-end-point of $\bar{U}$ is a local cut point*[7] *of $\bar{U}$, and* (c) *$\bar{U} - U$ is at most countable, and if T is locally non-separating at all points of $\bar{U}$, then $T(\bar{U})$ is a point.*

Proof. Assume $T(\bar{U})$ contains at least two points, y_1 and y_2. Consider any two points of $\bar{U}$, x_1 and x_2, such that $T(x_1) = y_1$ and $T(x_2) = y_2$. Then, since $\bar{U}$ is Peanian, $\bar{U}$ contains an arc α from x_1 to x_2. There are uncountably many points in $U \cdot \alpha$, because α is uncountable and $\bar{U} - U$ is countable. At most two points of $U \cdot \alpha$ can be end points of $\bar{U}$: hence $U \cdot \alpha$ contains uncountably many local cut points of $\bar{U}$. No point of $U \cdot \alpha$ which is a local cut point of $\bar{U}$ can be a component of any set of the form $\bar{U} \cdot T^{-1}(x)$, because if it were, then T would not be locally non-separating at that point.

It is not possible for all the points of $U \cdot \alpha$ which are local cut points of $\bar{U}$ to lie in one set $T^{-1}(x)$, for any closed set containing all such points contains all points of α, and then we would have $T(x_1) = T(x_2)$. This is a contradiction to our original assumption.

We now have that every point of $U \cdot \alpha$ which is a local cut point of $\bar{U}$ must lie in a non-degenerate component of a set of the form $\bar{U} \cdot T^{-1}(x)$, and not all such points can lie in one set $T^{-1}(x)$. The set $T(\alpha)$ is a connected set containing more than one point, and hence it contains uncountably many points. Therefore α has points in uncountably many sets of the type $\bar{U} \cdot T^{-1}(x)$, and all but a countable number of these points are points of U which are local cut points of $\bar{U}$. Therefore there are uncountably many mutually exclusive non-degenerate components of different sets $\bar{U} \cdot T^{-1}(x)$, each component having at least one point in common with $U \cdot \alpha$.

Of this set of components just described, only a countable number can be contained in the arc α. Therefore uncountably many of them must contain points whose distance from α is greater than some $\epsilon > 0$. Then it follows readily that in each of these components there is an arc which contains a point

[7] DEFINITION. A point p of a continuum M is called a *local cut point* of M provided there is a neighborhood U of p such that p is a cut point of the component of $\bar{U} \cdot M$ which contains p.

whose distance from α is greater than ϵ and which has exactly one point in common with α. The set M of points of intersection of these arcs with α is uncountable, and hence M contains a set of condensation points. Then there exists in M a point p_0 which is the limit point of a convergent sequence $\{p_i\}$ of points of M. Let α_i denote the arc which contains p_i. Then $\{\alpha_i\}$ $(i = 1, 2, \cdots)$ is a sequence of sets which converges to a continuum C_0 containing p_0. The decomposition of $\overline{U}$ into sets of the form $\overline{U} \cdot T^{-1}(x)$ is upper semi-continuous and since each set α_i is contained in a different set $\overline{U} \cdot T^{-1}(x)$, it follows that C_0 is a subcontinuum of the set $\overline{U} \cdot T^{-1}(x)$ which contains p_0. Then $C_0 \cdot \alpha_i = 0$ for $i = 1, 2, \cdots$. Furthermore, C_0 contains points whose distance from α is equal to or greater than ϵ, since each set α_i contains such points. Let A denote the minimal subarc of α containing all points p_i. Then let $K = A + C_0 + \sum_{i=1}^{\infty} \alpha_i$. The set K is connected and closed, and hence is a continuum. However, K is non-Peanian because it fails to be locally connected at any point of $C_0 - A$. This follows from the fact that in order to join a point of α_i to a point of $C_0 - A$ by a connected set in K it is necessary to pass through the point p_i in A. Thus we have in $\overline{U}$ a subcontinuum K which is non-Peanian. But it follows from the hypotheses of the theorem that $\overline{U}$ is hereditarily locally connected (cf. [8]). This is a contradiction. Hence it follows that $T(\overline{U})$ must be a single point, and the theorem is proved.

Corollary 5.11. *If A is a dendrite or finite graph and T is locally non-separating, then $T(A)$ is a point.*

Corollary 5.12. *If T is locally non-separating and A contains a connected open set U such that $\overline{U}$ is a finite graph, then $T(\overline{U})$ is a point.*

Corollary 5.13. *If every cyclic element of the Peano continuum A is a finite graph, and T is locally non-separating, then $T(A)$ is a point.*

The condition that $\overline{U} - U$ shall be at most countable, in the hypothesis of Theorem 5.1, is necessary, as the following example demonstrates: We start with a square plus its interior. On one edge of the square designate by K a Cantor non-dense perfect set. Then, by using the complementary intervals of this Cantor set as diameters, construct semi-circular loops outside the square. Denote by $\overline{U}$ the Cantor set K plus the semi-circular loops. Let $U = \overline{U} - K$. Then $\overline{U} - U = K$ is uncountably infinite. Decompose the entire set A (square plus loops) as follows: Let each loop plus the complementary interval on which it was constructed be an element and every other point be individually an element. This decomposition is upper semi-continuous and thus induces a continuous transformation T carrying each element into a point of a space B. This transformation is easily seen to be locally non-separating on the entire space A. Further, $\overline{U}$ satisfies all the conditions of Theorem 5.1 except that $\overline{U} - U$ is not countable. But $T(\overline{U})$ is not a point: in fact it is an arc of the space B. Hence the condition that $\overline{U} - U$ be countable is necessary.

6. **Locally non-alternating transformations.** The definition of a locally non-alternating transformation, as given by Vance, is as follows: A continuous transformation is called *locally non-alternating at a point* p of $T^{-1}(x)$ provided for any neighborhood U_p of p, there exists a neighborhood V_p lying in U_p such that if y is any point of B distinct from x, and u and v are any two points of $T^{-1}(y) \cdot V_p$, then u and v lie in a connected subset of $U_p - T^{-1}(x)$. A continuous transformation is called *locally non-alternating on a space* A if it is locally non-alternating at all points of A.

Vance gave simple examples to show that a non-alternating transformation is not necessarily locally non-alternating, and that a locally non-alternating transformation is not necessarily non-alternating.

THEOREM 6.1. *If T is a light*[8] *locally non-alternating transformation from the dendrite A to the dendrite B, then T is a homeomorphism.*

Proof. Assume T is not a homeomorphism. Then some two points a_0 and b_0 of A must go into the same point x_0 of B. Consider the arc $a_0b_0 = \alpha_0$. It must contain a point p such that $T(p) = q \neq x_0$, for otherwise $T^{-1}(x_0)$ would contain the connected set α_0 and hence T would not be light. Also, there must be some point x_1 of B, distinct from q and x_0, such that $T^{-1}(x_1) \cdot \text{arc } a_0p \neq 0$ and $T^{-1}(x_1) \cdot \text{arc } pb_0 \neq 0$, for if there were no such point x_1, then we would have $T(\text{arc } a_0p) \cdot T(\text{arc } pb_0) = q + x_0$. Thus there would be two connected sets running from q to x_0 and having only $q + x_0$ in common, but this is impossible in a dendrite. Let a_1 denote the first point of $T^{-1}(x_1)$ in the order from p to a_0; b_1 the first point of $T^{-1}(x_1)$ in the order from p to b_0. Then arc $a_1b_1 = \alpha_1$ lies entirely in α_0. By repeating the above argument we can obtain an arc α_2 lying within α_1, an arc α_3 within α_2, etc. Repeat the process indefinitely.

Let $\alpha_\omega = \prod_{i=1}^{\infty} \alpha_i$. If α_ω is a point, then T is not locally non-alternating at this point because $a_i \rightarrow \alpha_\omega$, $b_i \rightarrow \alpha_\omega$ and $T(a_i) = T(b_i)$, but a_i and b_i cannot be connected inside small neighborhoods of α_ω except through α_ω. Since α_ω is not a point, it must be an arc with end points a_ω and b_ω, and it follows from the continuity of T that $T(a_\omega) = T(b_\omega)$. Then the arc $\alpha_{\omega+1}$ can be obtained just as the arcs α_i $(i < \omega)$ were obtained. This process may be continued indefinitely and arcs α_β defined for all ordinals of the first and second classes, unless for some limiting ordinal β, α_β is a point. But this is impossible for at this point T would fail to be locally non-alternating. But the process cannot continue through all the ordinals of the first and second classes, for then A would contain an uncountable monotonic decreasing sequence of continua, and this is impossible.

Hence T must be one-to-one, and therefore a homeomorphism.

[8] DEFINITION. A transformation T is said to be light if every set $T^{-1}(x)$ is totally disconnected.

The following theorems proved by Vance follow as corollaries of this theorem:

COROLLARY 6.11. *If T is locally non-alternating, with A an arc and B a dendrite, and if $T^{-1}(x)$ contains no open set, then T is a homeomorphism and thus B is an arc.*

COROLLARY 6.12. *If T is locally non-alternating where B is a dendrite and A is a dendrite in which the branch points are dense on no arc of A, and if $T^{-1}(x)$ contains no open set, then T is a homeomorphism.*

The remainder of this section will be devoted to consideration of the relationship between locally non-alternating transformations and 0-regular transformations.[9]

THEOREM 6.2. *If the transformation $T(A) = B$ is 0-regular, then T is locally non-alternating on A.*

Proof. This theorem follows readily with the help of the following theorem of Wallace [6]: If A is a compact continuum, a necessary and sufficient condition that the interior transformation $T(A) = B$ shall be 0-regular is that for each $\epsilon > 0$ there exist a $\delta > 0$ such that if u and v are in A with $\rho(u, v) < \delta$ and $T(u) = T(v)$, then u and v lie in an ϵ-continuum in $T^{-1}[T(u)] = T^{-1}[T(v)]$. Let p be any point of A, and let U_p be any neighborhood of p. Then, since we are dealing with a metric space, U_p contains a spherical neighborhood S_p with center at p and radius equal to k, for some k. Let $\epsilon = \frac{1}{2}k$. Then, for this ϵ, there is determined a number $\delta > 0$ which satisfies the condition given in the theorem of Wallace just quoted. Let V_p be a spherical neighborhood with center at p and radius equal to $\frac{1}{2}\delta$. Now consider any point y distinct from $T(p)$ in B, and let u and v be any two points of $T^{-1}(y) \cdot V_p$. Then $\rho(u, v) < \delta$, and hence u and v lie in an ϵ-continuum C in $T^{-1}(y)$, by Wallace's theorem. Because $\epsilon = \frac{1}{2}k$, C is contained in S_p and hence in U_p. Furthermore, since $y \neq T(p)$ and C is contained in $T^{-1}(y)$, $C \cdot T^{-1}[T(p)] = 0$. Hence u and v, any two points of $T^{-1}(y) \cdot V_p$, lie in a connected subset of $U_p - T^{-1}[T(p)]$. Therefore, by definition, T is locally non-alternating at the point p. Since p was any point of A, T is locally non-alternating on A.

The converse of Theorem 6.2 is not true in general, i.e., an interior, locally non-alternating transformation is not necessarily 0-regular. The following example demonstrates this remark: Let A denote a spherical shell of outer radius unity and inner radius $\frac{1}{2}$. Let T be the orthogonal projection of A onto a tangent plane. This transformation is interior and locally non-alternating

[9] DEFINITION. (1) If the sequence of closed sets M_n converges to M, then $M_n \to M$ 0-*regularly* provided that for each $\epsilon > 0$ there are positive numbers δ and N such that for $n > N$ any two points x and y in M_n with $\rho(x, y) < \delta$ lie in a continuum in M_n of diameter less than ϵ.

(2) A continuous transformation T is called 0-*regular* provided that if $y_n \to y$ in B, the sets $T^{-1}(y_n) \to T^{-1}(y)$ 0-regularly in A.

at all points of A, but it fails to be 0-regular at any point of the inner surface of A whose distance from the tangent plane is unity.

By looking carefully at the requirements of 0-regular and locally non-alternating transformations it is easy to see why 0-regularity is the stronger type of transformation. In non-rigorous terms, 0-regularity requires that if u and v, two points of a set $T^{-1}(y)$, are sufficiently close together, they must lie in a small connected set C *which is also contained in* $T^{-1}(y)$. Locally non-alternating merely requires that two such points u and v shall lie in a small connected set C, and does *not* require that C shall be contained in $T^{-1}(y)$, but only that it shall avoid $T^{-1}(x)$ for a specified point x. This is obviously a weaker condition, and so the locally non-alternating condition is not sufficient in general to imply 0-regularity.

THEOREM 6.3. *On a dendrite, 0-regular transformations and interior, locally non-alternating transformations are equivalent, since both are homeomorphisms.*

Proof. It was proved by Wallace [6] that any 0-regular transformation on a dendrite is a homeomorphism. G. T. Whyburn showed that on a dendrite any interior transformation is also light (cf. [11]), and that the image of a dendrite under an interior transformation is again a dendrite cf. [10]. Then by Theorem 6.1 at the beginning of this section, it follows that an interior locally non-alternating transformation on a dendrite is a homeomorphism. Hence the theorem follows.

BIBLIOGRAPHY

1. B. KNASTER AND C. KURATOWSKI, *Sur les ensembles connexes*, Fundamenta Mathematicae, vol. 2(1921), p. 214.
2. C. KURATOWSKI, *Les décompositions semi-continues*, Fundamenta Mathematicae, vol. 11(1928), pp. 169–185.
3. C. KURATOWSKI, *Sur les continus de Jordan*, Fundamenta Mathematicae, vol. 5(1924), p. 119.
4. R. L. MOORE, *Concerning upper semi-continuous collections of continua*, Transactions of the American Mathematical Society, vol. 27(1925), pp. 416–428.
5. E. P. VANCE, *Generalizations of non-alternating and non-separating transformations*, this Journal, vol. 6(1940), pp. 66–79.
6. A. D. WALLACE, *On 0-regular transformations*, American Journal of Mathematics, vol. 62(1940), pp. 277–284.
7. J. F. WARDWELL, *Non-separating transformations*, this Journal, vol. 2(1936), pp. 745–750.
8. G. T. WHYBURN, *Concerning points of continuous curves defined by certain im kleinen properties*, Mathematische Annalen, vol. 102(1930), p. 321, Theorem 11.
9. G. T. WHYBURN, *Non-alternating transformations*, American Journal of Mathematics, vol. 56(1934), pp. 294–302.
10. G. T. WHYBURN, *Interior transformations on compact sets*, this Journal, vol. 3(1937), pp. 370–381.
11. G. T. WHYBURN, *Interior transformations on certain curves*, this Journal, vol. 4(1938), pp. 607–612.

UNIVERSITY OF MICHIGAN.

A NEW PROOF OF A THEOREM OF MENCHOFF

By Raphaël Salem

1. Let $f(x)$ be a function of integrable square, of period 2π, and let $S_n(x)$ be the sum of order n of its Fourier series. It has been shown in a previous paper[1] that the study of the integral

$$\int_0^{2\pi} S_{n(x)}(x)\,dx \qquad (n(x) \leqq n)$$

considered by Kolmogoroff and Seliverstoff for the study of the almost-everywhere convergence of the series is equivalent to the study of the sum

$$T_n = \sum_1^n \left[\left(\int_0^{2\pi} f_p \cos px\,dx \right)^2 + \left(\int_0^{2\pi} f_p \sin px\,dx \right)^2 \right],$$

$\{f_p(x)\}$ being a sequence of characteristic functions of sets such that $f_p \geqq f_{p+1}$; it has been proved in the same paper that $T_n < C \log n$, C being an absolute constant.[2]

The purpose of this paper is to consider sums of the type

$$\sum_1^n \left(\int_0^1 f_p \varphi_p\,dx \right)^2,$$

$\{f_p(x)\}$ being a decreasing sequence of characteristic functions of sets, and $\{\varphi_p(x)\}$ any family of orthogonal normal functions in (0, 1). The study of such sums will be carried out as an application of Bessel's inequality in a two-dimensional domain.

2. Let $p_1, p_2, \cdots, p_n$ and P be functions of a certain number of variables $x_1, x_2, \cdots, x_k$ defined in a k-dimensional domain D, and let us write, for the sake of brevity, $\int Q\,d\tau$ instead of $\int_D Q(x_1, \cdots, x_k)\,dx_1 \cdots dx_k$. Then, assuming that

$$\int p_i p_j\,d\tau = 0 \qquad (i \neq j), \tag{1}$$

Received September 21, 1940.

[1] See Comptes Rendus Acad. Sci. Paris, vol. 205(1937), pp. 14–16. See also Salem, *Essais sur les Séries Trigonométriques*, Paris, 1940.

[2] In this paper C means any absolute constant, not necessarily the same throughout the paper.

we see that Bessel's inequality is

$$\sum_1^n \frac{\left(\int P p_k \, d\tau\right)^2}{\left(\int p_k^2 \, d\tau\right)} \leqq \int P^2 \, d\tau.$$

Let us now apply this general inequality in the case of two variables x, t, in the square $0 \leqq x \leqq 1$, $0 \leqq t \leqq 1$. Let us put

$$P(x, t) = \sum_1^n f_k(x) g_k(t); \qquad p_k(x, t) = \varphi_k(x) h_k(t).$$

$\{f_k\}$ and $\{\varphi_k\}$ being defined as at the end of §1, the equalities (1) hold good. Let us suppose now

$$\int_0^1 g_i(t) h_j(t) \, dt = 0 \qquad (i \neq j); \tag{2}$$

then

$$\int P p_k \, d\tau = \int_0^1 f_k(x) \varphi_k(x) \, dx \cdot \int_0^1 g_k(t) h_k(t) \, dt$$

and Bessel's inequality becomes

$$\sum_1^n \left(\int_0^1 f_k \varphi_k \, dx\right)^2 \frac{\left(\int_0^1 g_k h_k \, dt\right)^2}{\int_0^1 h_k^2 \, dt} \leqq \int_0^1 \int_0^1 \left(\sum_1^n f_k(x) g_k(t)\right)^2 dx \, dt, \tag{3}$$

an inequality which is true on the sole assumptions that $\{\varphi_k\}$ be an orthogonal normal system, and that the equalities (2) be satisfied.

Let us now introduce the hypothesis that $\{f_k\}$ is a decreasing sequence of characteristic functions of sets. We get

$$\sum_1^n f_k(x) g_k(t) = \sum_1^n \Delta f_k(x) \cdot \sigma_k(t),$$

putting

$$\Delta f_k = f_k - f_{k+1}, \qquad \Delta f_n = f_n, \qquad \sigma_k = g_1 + \cdots + g_k.$$

Then

$$\left(\sum_1^n f_k(x) g_k(t)\right)^2 = \sum_1^n \Delta f_k(x) \sigma_k^2(t).$$

Hence

$$\int_0^1 \int_0^1 \left(\sum_1^n f_k(x) g_k(t)\right)^2 dx \, dt \leqq \max_k \int_0^1 \sigma_k^2(t) \, dt,$$

and we get the inequality

$$(4) \qquad \sum_1^n \left(\int_0^1 f_k \varphi_k \, dx\right)^2 \frac{\left(\int_0^1 g_k h_k \, dt\right)^2}{\int_0^1 h_k^2 \, dt} \leqq \max_k \int_0^1 [g_1(t) + \cdots + g_k(t)]^2 \, dt$$

in the right side of which it is interesting to observe that we have only to take into account the values of k for which $\Delta f_k \neq 0$.

3. Let us apply this inequality in the following particular case:

$$g_k(t) = \sqrt{t} \cdot \sin 2\pi kt, \qquad h_k(t) = \frac{\sin 2\pi kt}{\sqrt{t}}.$$

The equalities (2) hold good and we have

$$\int_0^1 h_k^2(t) \, dt < A \log k, \qquad \int_0^1 [g_1 + \cdots + g_k]^2 \, dt < B \log k,$$

A, B being absolute constants. Hence, the inequality (4) gives

$$\sum_1^n \frac{1}{\log k} \left(\int_0^1 f_k \varphi_k \, dx\right)^2 < C \log n$$

or

$$(5) \qquad \sum_1^n \left(\int_0^1 f_k \varphi_k \, dx\right)^2 < C \log^2 n,$$

C being an absolute constant.

4. It has been shown in the papers quoted in §1 that this last inequality leads immediately to the conclusion that if $f(x)$ is of integrable square and S_n is the sum of order n of the orthogonal series $\sum c_p \varphi_p$ representing f, we have

$$\int_0^1 S_{n(x)}(x) \, dx < C \log n \sqrt{\sum_1^n c_p^2} \qquad (n(x) \leqq n).$$

Furthermore, there is little to change in the argument of §2 and §3 in order to get the stronger inequality

$$(6) \qquad \int_0^1 S_{n(x)}^2(x) \, dx < C \log^2 n \sum_1^n c_p^2 \qquad (n(x) \leqq n).$$

For if we multiply every function f_k in (3) by any function of integrable square $F(x)$, the inequality (3) holds good, and the inequality (4) becomes easily

$$(4') \qquad \sum_1^n \left(\int_0^1 F \cdot f_k \cdot \varphi_k \, dx\right)^2 \frac{\left(\int_0^1 g_k h_k \, dt\right)^2}{\int_0^1 h_k^2 \, dt} \leqq \max_k \int_0^1 F^2 \, dx \cdot \int_0^1 [g_1 + \cdots + g_k]^2 \, dt$$

and (5) becomes

$$(5') \qquad \sum_1^n \left(\int_0^1 F \cdot f_k \cdot \varphi_k \, dx \right)^2 < C \log^2 n \int_0^1 F^2(x) \, dx.$$

We get immediately the inequality

$$\int_0^1 F(x) S_{n(x)}(x) \, dx < C \log n \sqrt{\sum_1^n c_p^2} \left(\int_0^1 F^2(x) \, dx \right)^{\frac{1}{2}}$$

for every function $F(x)$ of integrable square, and this is equivalent to (6). I am indebted for this remark to Professor Zygmund.

5. This proof of inequality (6) gives a new proof of the well-known theorem of Menchoff on the convergence almost-everywhere of the series $\sum c_p \varphi_p$ if the series $\sum c_n^2 \log^2 n$ converges.

It is known that this theorem cannot be improved if the orthogonal sequence $\{\varphi_n\}$ is arbitrary. This leads to an interesting conclusion. Let $\{\theta_k(t)\}$ be an orthogonal, normal, uniformly bounded set of functions in (0, 1) and let $\{c_k\}$ be any sequence of coefficients such that $| c_k | \log k$ tends to infinity with k. Then the absolute values of the partial sums of the series $\sum c_k \theta_k(t)$ can never be majorized by an integrable function $\psi(t)$.

For let us put in the inequality (4)

$$g_k(t) = \frac{c_k \theta_k(t)}{\sqrt{\psi(t)}}, \qquad h_k(t) = \theta_k(t) \sqrt{\psi(t)}.$$

We get, assuming that $\int_0^1 \psi(t) \, dt < A$ and $| \theta_k(t) | < M$,

$$\sum_1^n \left(\int_0^1 f_k \varphi_k \, dx \right)^2 \frac{c_k^2}{M^2 A} < \max_k \int_0^1 \frac{1}{\psi(t)} \left(\sum_1^k c_k \theta_k(t) \right)^2 dt.$$

Then, if we suppose

$$\left| \sum_1^k c_k \theta_k(t) \right| < \psi(t),$$

we get the inequality

$$\sum_1^n \left(\int_0^1 f_k \varphi_k \, dx \right)^2 c_k^2 < A^2 M^2$$

which is impossible in the hypothesis $| c_k | \log k \to \infty$, for otherwise Menchoff's theorem on almost-everywhere convergence would not be the best one.

Thus, for instance, for no sequence of integers $\{n_k\}$ can the absolute values of the partial sums of the series

$$\sum_1^\infty r_k \cos (n_k x - \alpha_k)$$

be majorized by a summable function if $r_k \log k$ tends to infinity.

MONTREAL, CANADA.

A TOPOLOGY FOR SEMI-MODULAR LATTICES

BY L. R. WILCOX

Introduction

In classical differential geometry of Euclidean, affine and projective spaces, essential use is made of an underlying topology of the point space. Moreover, there is needed in these theories an extension of this topology to the lattice of linear subsets, for the purpose, for example, of defining tangent lines and planes, osculating planes, etc.; such an extension is generally not considered explicitly but is usually tacitly assumed. In view of the recent developments[1] in the theory of lattices, particularly those which relate lattices to geometries, it seems desirable to consider the question of constructing such an extension abstractly and in a lattice-theoretic manner, sufficiently generally to include the topological foundations of the familiar differential geometries. The present paper gives a solution of this problem for a *semi-modular* lattice[2] and considers several important questions suggested by our investigation; further questions, some of which are mentioned at the close of the paper, are reserved for a subsequent publication.

Inasmuch as we shall consider throughout the entire paper a fixed semi-modular lattice L which is *atomistic* and satisfies *chain conditions*,[3] we devote the first section to a statement of our axioms and to some immediate consequences and fundamental lemmas. A topology for the set P of *points* (or atoms) of L is introduced by assuming P to be a metric space. In the second section, the topology of P is extended to L in a natural way; it is shown that L is a topological space in the sense of Alexandroff and Hopf.[4] §3 deals with the introduction of a complete system of neighborhoods in L and a proof that L is a Hausdorff space. The paper closes with a brief study of the continuity of the lattice operations.

Frequent reference will be made to a paper by the author[5] (to be referred to as M. T. L.), wherein the notions of *modularity*, *independence* and *semi-modular lattice* are considered in detail; the definitions and results of M. T. L. essential to the present work will be assumed, but for completeness these will be stated

Received October 7, 1940; presented to the American Mathematical Society, September 10, 1940.

[1] See G. Birkhoff, *Lattice Theory*, New York, 1940, particularly Chapter IV.

[2] See §1 for the definition of modularity.

[3] See §1 for the definition of atomistic lattice. For a statement of the chain conditions, see L. R. Wilcox, *Modularity in the theory of lattices*, Annals of Math., vol. 40(1939), p. 497.

[4] P. Alexandroff and H. Hopf, *Topologie*, Berlin, 1935, p. 37.

[5] L. R. Wilcox, *Modularity in the theory of lattices*, Annals of Math., vol. 40(1939), pp. 490–505.

in §1. All notations here are the same as were used in M. T. L. Positive real numbers are always denoted by ϵ, η; integers are denoted by k, l, m, n, while μ, ν are used for indices in infinite sequences.

1. **Foundations of the theory.** Let L be a lattice satisfying both ascending and descending chain conditions,[6] so that $0 \in L$ and $1 \in L$ exist, and so that every set $S \subset L$ has a *sum* (least upper bound) and a *product* (greatest lower bound),[7] to be denoted by $\sum S$ and $\prod S$ respectively. By $(b, c)M$ (read b and c are *modular*) we mean

$$a \leqq c \textit{ implies } (a + b)c = a + bc.$$

By $(b, c)\perp$ (read b and c are *independent*) we mean $bc = 0$ and $(b, c)M$. By $(a_1, \cdots, a_n)\perp$ is meant $(\sum (a_i, i \in S), \sum (a_j; j \in T))\perp$ for $S, T \subset [1, \cdots, n]$ such that $i \in S$, $j \in T$ implies $i < j$. A necessary and sufficient condition for $(a_i)\perp$ is[8] $(a_i, a_{i+1} + \cdots + a_n)\perp$ for $i = 1, \cdots, n - 1$. We say that L is *semi-modular* in case

(a) $(a, b)\perp$ *implies* $(b, a)M$;

(b) $ab \neq 0$ *implies* $(a, b)M$.

By $a \prec b$ (or $b \succ a$) we mean $a < b$ together with $a \leqq c \leqq b$ implies $c = a$ or $c = b$. A *point* is an element $p \in L$ such that $p \succ 0$; the set of all points is denoted by P, and points are denoted generically by p, q, r. For $a \in L$ we define $P_a \equiv [p; p \leqq a]$. The lattice L is called *atomistic* in case

$$a \in L \textit{ implies } a = \sum P_a.$$

If $a \leqq b$, we say that c is a *complement* of a in b in case $a + c = b$, $ac = 0$; c is a *modular complement* in case $(a, c)\perp$. A *basis* of $a \in L$ is a finite system $(p_1, \cdots, p_n)$ where $p_i \in P$, $p_1 + \cdots + p_n = a$ and $(p_1, \cdots, p_n)\perp$.

In what follows we shall assume

AXIOM I. *L is semi-modular, satisfies the ascending and descending chain conditions, and is atomistic.*

It is proved in M. T. L. that there exist an integer $N = 0, 1, 2, \cdots$ and a (unique) "dimension function" $(d(a); a \in L)$ which has the properties

(a) $d(0) = 0$, $d(1) = N$;

(b) $[d(a); a \in L] = [0, 1, \cdots, N]$;

(c) $d(a) < d(b)$ *for* $a < b$;

(d) $d(a) + 1 = d(b)$ *for* $a \prec b$;

(e) $d(a + b) + d(ab) \leqq d(a) + d(b)$;

(f) $d(a + b) + d(ab) = d(a) + d(b)$ *if and only if* $(a, b)M$.

Extensive use will be made of this dimension function. Note that the nor-

[6] See M. T. L., pp. 497, 491.

[7] See L. R. Wilcox and M. F. Smiley, *Metric lattices*, Annals of Math., vol. 40(1939), p. 310.

[8] M. T. L., p. 493.

malization in (a) is slightly different from that in M. T. L.[9] It is also proved in M. T. L. that the relation $(a_1, \cdots, a_n)\perp$ is symmetric.[10]

Axiom II. *P is a metric space with respect to a metric $(\delta(p, q); p, q \in P)$.*

Definition 1.1. If $S, T \subset P$, and $S, T \neq \Theta$, then

$$\delta(S, T) \equiv \text{g.l.b. } [\delta(p, q); p \in S, q \in T].$$

If $p \in P$, $a \in L$, $a \neq 0$, then

$$\delta(p, a) \equiv \delta(a, p) \equiv \delta([p], P_a).$$

Axiom III. *Let $p \in P$, $(p_1, \cdots, p_n)\perp$, $p_i \in P$. Then for every $\epsilon > 0$ there exists $\eta > 0$ such that*

$$| \delta(p, p_1 + \cdots + p_n) - \delta(p, q_1 + \cdots + q_n) | < \epsilon$$

for $\delta(q_i, p_i) < \eta$.

Remark. Axiom III asserts the continuity of $\delta(p, p_1 + \cdots + p_n)$ in the p_i at every place where $(p_i)\perp$.

Axiom IV. *The sets P_a are closed in the topology of P.*

Corollary. *$\delta(p, a) = 0$ if and only if $p \leqq a$.*

Proof. If $\delta(p, a) = 0$, there exists a sequence $p_\mu \leqq a$ such that $\lim \delta(p_\mu, p) = 0$. Hence p is in the closure of P_a, i.e., in P_a by Axiom IV; thus $p \leqq a$. The other implication is trivial.

Remark. It may be readily verified that our Axioms I–IV are satisfied in real or complex projective and Euclidean spaces (and hence also in affine spaces); in the projective cases δ is the elliptic metric, while in the others δ is the ordinary Euclidean distance. In all cases our relation $(p_1, \cdots, p_n)\perp$ is linear independence.

A few consequences of Axiom I not proved in M. T. L. will be given at this point.

Lemma 1.1. *If $(a_1, \cdots, a_n)\perp$, then*

$$d(a_1 + \cdots + a_n) = d(a_1) + \cdots + d(a_n),$$

and conversely.

Proof. This is obvious for $n = 1$. If it holds for $n = k$, then by (a), (f)

$$\begin{aligned} d(a_1 + \cdots + a_k + a_{k+1}) &= d(a_1 + \cdots + a_k) + d(a_{k+1}) \\ &= d(a_1) + \cdots + d(a_{k+1}), \end{aligned}$$

and the statement is true for $n = k + 1$. The converse is established by an inductive proof that $(a_1 + \cdots + a_k, a_{k+1})\perp$ for $k = 1, \cdots, n - 1$.

[9] M. T. L., p. 505.
[10] M. T. L., p. 496.

THEOREM 1.1. *If* $a \leqq b$, *then there exists* $c \in L$ *such that* $a + c = b$, $(a, c)\perp$, *and such that* c *is of the form* $p_1 + \cdots + p_k$, *with* $p_i \in P$, $(p_1, \cdots, p_k)\perp$ $(k = 0, 1, \cdots, N)$.

Proof. Let $d(a) = m$, $d(b) = l$. If $m = l$, then $c \equiv 0$ is effective. Suppose $m < l$. Then $a < b$, and, by the atomistic property of L, $P_a \subset P_b$, $P_a \neq P_b$. Let $p \in P_b - P_a$. Now $0 \leqq pa \leqq p$, whence $pa = p$ or $pa = 0$. If $pa = p$, then $p \leqq a$, contrary to[11] $p \in' P_a$. Hence $pa = 0$, and[12] $(p, a)\perp$. Thus there exists $p_1 \leqq b$ with $(p_1, a)\perp$, $p_1 + a \leqq b$. Inductive application of this argument yields a finite or infinite sequence $p_1, p_2, \cdots \leqq b$ with $(a + p_1 + \cdots + p_{i-1}, p_i)\perp$. If $b_i \equiv a + \cdots + p_{i-1}$, then $b_i \leqq b_{i+1}$; but $b_i = b_{i+1}$ implies $p_i \leqq a + p_1 + \cdots + p_{i-1}$, and this is impossible. Hence $b_i < b_{i+1}$, and our sequence (p_i) must be finite and of the form $(p_1, \cdots, p_k)$. Then if $c \equiv p_1 + \cdots + p_k$, we have $a + c = b$; since $(a + p_1 + \cdots + p_{i-1}, p_i)\perp$, it follows that $(a, p_1, \cdots, p_k)\perp$, whence $(a, c)\perp$. The statement that $k \leqq N$ follows from Lemma 1.1, since $k = d(c) \leqq d(1) = N$.

Remark. Theorem 1.1 states that every a which $\leqq b$ has a modular complement in b. It should be observed that this proposition is equivalent to the statement that L is atomistic, if the chain conditions are assumed.

THEOREM 1.2. *Suppose* $a \in L$, $a \neq 0$, $d(a) = k$. *Then there exists a basis* $(p_1, \cdots, p_k)$ *of* a, *and every basis of* a *consists of* k *points.*

Proof. Applying Theorem 1.1 to $0 \leqq a$, we obtain $c = p_1 + \cdots + p_m$ such that $0 + c = a$. Hence the points p_i are a basis of a. The second part and the fact that $m = k$ follow from Lemma 1.1.

Remark. The results of Theorem 1.2 will be used freely without reference in what follows. They state the property which is usually used in geometry to define dimensionality.

THEOREM 1.3. *If* $S \subset P$, $S \neq \Theta$, *then there exist* $k = 1, \cdots, N$ *and* $p_1, \cdots, p_k \in S$ *such that* $\sum S = p_1 + \cdots + p_k$, $(p_i)\perp$.

Proof. Since $\sum p_i \leqq \sum S$ for $p_i \in S$, it suffices to prove that $\sum p_i < \sum S$ cannot occur for every choice of $p_1, \cdots, p_k$ with $p_i \in S$, $(p_i)\perp$. Let us suppose the contrary. Then for every such system $(p_1, \cdots, p_k)$ there exists $p_{k+1} \in S$ with $p_1 + \cdots + p_{k+1} < \sum S$, $(p_1, \cdots, p_{k+1})\perp$; for otherwise $p \in S$ implies $p \leqq p_1 + \cdots + p_k$, i.e., $\sum S = p_1 + \cdots + p_k$, contrary to the hypothesis. Since $S \neq \Theta$, inductive definition yields an infinite sequence $(p_\mu; \mu = 1, 2, \cdots)$ such that $(p_1, \cdots, p_n)\perp$ for $n = 1, 2, \cdots$. Then if $a_\mu \equiv p_1 + \cdots + p_\mu$, the sequence (a_μ) has the property $a_\mu < a_{\mu+1}$, contrary to the ascending chain condition. Obviously $k \leqq N$. This completes the proof.

2. Extension to L of the topology of P.

[11] The prime is used as a symbol for negation; thus $p \in' P_a$ means p is not an element of P_a.

[12] M. T. L., p. 491.

LEMMA 2.1. *Let* $(a^\mu;\ \mu = 1, 2, \cdots)$, $a^\mu \neq 0$, *and* $p \in P$ *be given. If* $\lim \delta(p, a^\mu) = 0$, *then there exists* $(p^\mu; \mu = 1, 2, \cdots)$, $p^\mu \leqq a^\mu$, *such that* $\lim p^\mu = p$.

Proof. Clearly there exists a sequence (p^μ) with $p^\mu \leqq a^\mu$ such that

$$0 \leqq \delta(p, p^\mu) - \delta(p, a^\mu) < \frac{1}{\mu};$$

hence

$$0 \leqq \lim \delta(p, p^\mu) - \lim \delta(p, a^\mu) \leqq 0,$$

and it follows that $\lim \delta(p, p^\mu) = 0$.

LEMMA 2.2. *If* $(p_1, \cdots, p_k)\perp$, $\lim_\mu p_i^\mu = p_i$ *for* $i = 1, \cdots, k$, $p \leqq p_1 + \cdots + p_k$, *then* $\lim_\mu \delta(p, p_1^\mu + \cdots + p_k^\mu) = 0$.

Proof. Let $\epsilon > 0$ be given. Then by III there exists $\eta > 0$ such that

$$| \delta(p, p_1 + \cdots + p_k) - \delta(p, q_1 + \cdots + q_k) | < \epsilon$$

when $\delta(q_i, p_i) < \eta$. Since $\lim_\mu p_i^\mu = p_i$, there exists $m = 1, 2, \cdots$ such that $\delta(p_i^\mu, p_i) < \eta$ for $\mu > m$. Hence for $\mu > m$

$$| \delta(p, p_1 + \cdots + p_k) - \delta(p, p_1^\mu + \cdots + p_k^\mu) | < \epsilon.$$

Since $\delta(p, p_1 + \cdots + p_k) = 0$ by virtue of the hypothesis $p \leqq \sum p_i$ and the corollary to Axiom IV, the proof is complete.

LEMMA 2.3. *Suppose* $p, q \in P$, $a \in L$, $a \neq 0$. *Then*

$$| \delta(q, a) - \delta(p, a) | \leqq \delta(p, q).$$

Proof. For $r \in P$, $r \leqq a$, we have

$$\delta(q, r) \leqq \delta(p, r) + \delta(p, q),$$

whence

$$\text{g.l.b. } [\delta(q, r); r \leqq a] \leqq \text{g.l.b. } [\delta(p, r); r \leqq a] + \delta(p, q),$$

and $\delta(q, a) \leqq \delta(p, a) + \delta(p, q)$. Therefore $\delta(q, a) - \delta(p, a) \leqq \delta(p, q)$. By symmetry $\delta(p, a) - \delta(q, a) \leqq \delta(p, q)$. Thus the proof is complete.

LEMMA 2.4. *The function* $\delta(p, p_1 + \cdots + p_k)$ *is continuous in* $p, p_1, \cdots, p_k$ *at every place where* $(p_1, \cdots, p_k)\perp$.

Proof. Suppose $\epsilon > 0$. Then by Axiom III there exists $\eta_1 > 0$ such that

$$| \delta(p, p_1 + \cdots + p_k) - \delta(p, q_1 + \cdots + q_k) | < \tfrac{1}{2}\epsilon$$

when $\delta(q_i, p_i) < \eta_1$. Let η be the smaller of $\frac{1}{2}\epsilon$ and η_1. Then for $\delta(q, p)$, $\delta(q_i, p_i) < \eta$ we have, using Lemma 2.3,

$$\begin{aligned} | \delta(p, \textstyle\sum p_i) - \delta(q, \textstyle\sum q_i) | \\ &\leqq | \delta(p, \textstyle\sum p_i) - \delta(p, \textstyle\sum q_i) | + | \delta(p, \textstyle\sum q_i) - \delta(q, \textstyle\sum q_i) | \\ &\leqq \tfrac{1}{2}\epsilon + \delta(p, q) < \tfrac{1}{2}\epsilon + \tfrac{1}{2}\epsilon = \epsilon. \end{aligned}$$

Remark. Lemma 2.4 is a slight extension of the continuity assumed in Axiom III.

DEFINITION 2.1. For a sequence $(a^\mu; \mu = 1, 2, \cdots)$ we define $L(a^\mu) = L(a^\mu; \mu = 1, 2, \cdots)$ as the set of all $p \in P$ such that there exist $n = 1, 2, \cdots$ and a sequence $(p^\mu; \mu = n + 1, n + 2, \cdots)$ with $p^\mu \leqq a^\mu$ $(\mu > n)$ and $\lim_{\mu > n} p^\mu = p$.

COROLLARY. *If $a^\mu \in L$ for $\mu = 1, 2, \cdots$, then for every $n = 1, 2, \cdots$*

$$L(a^\mu; \mu = 1, 2, \cdots) = L(a^\mu; \mu = n + 1, n + 2, \cdots).$$

The proof is obvious.

THEOREM 2.1. *If $a^\mu \in L$ for $\mu = 1, 2, \cdots$, then there exists a unique $a \in L$ such that $L(a^\mu) = P_a$; and $a = \sum L(a^\mu)$.*

Proof. The uniqueness of a is obvious from the atomistic property of L, since $P_a = P_b$ implies $a = b$. Hence it remains to show only that $L(a^\mu) = P_a$ for $a \equiv \sum L(a^\mu)$. Suppose $p \in L(a^\mu)$. Then $p \leqq a$, i.e., $p \in P_a$. Thus $L(a^\mu) \subset P_a$. Suppose now that $p \in P_a$, i.e., $p \leqq a$. Then $L(a^\mu) \neq \Theta$, since otherwise $a = 0$, and this is impossible. By Theorem 1.3 there exist $k = 1, 2, \cdots$ and $p_1, \cdots, p_k \in L(a^\mu)$ such that $\sum p_i = a$, $(p_i)\perp$. Hence there exist $n = 1, 2, \cdots$ and sequences $p_i^\mu \leqq a^\mu$ $(\mu > n)$ such that $p_i = \lim_{\mu > n} p_i^\mu$ for $i = 1, \cdots, k$. Since $p \leqq \sum p_i$, we may apply Lemma 2.2 and conclude that

$$\lim_{\mu > n} \delta(p, \sum [p_i^\mu; i = 1, \cdots, k]) = 0.$$

Now by Lemma 2.1 there exists a sequence $(p^\mu; \mu > n)$ such that $p^\mu \leqq \sum_i p_i^\mu \leqq a^\mu$, $\lim p^\mu = p$. Thus $p \in L(a^\mu)$, and $P_a \subset L(a^\mu)$. This completes the proof.

COROLLARY. *If $a, a^\mu \in L$ $(\mu = 1, 2, \cdots)$, then $a = \sum L(a^\mu)$ if and only if the statements* (a) *$p \leqq a$ and* (b) *p is of the form $\lim_{\mu > n} p^\mu$ with $p^\mu \leqq a^\mu$ $(\mu > n)$ are equivalent.*

The proof is obvious.

DEFINITION 2.2. If $a^\mu \in L$ $(\mu = 1, 2, \cdots)$, then (a^μ) is convergent in case

$$\lim d(a^\mu) = d(\sum L(a^\mu)).$$

When (a^μ) is convergent, we write $\lim a^\mu$ for $\sum L(a^\mu)$. (The statement $a = \lim a^\mu$ always means that (a^μ) is convergent and that $a = \sum L(a^\mu)$.)

Remark. It may be verified that if $a^\mu \in P$ the statement $a = \lim a^\mu$ is equivalent to $a \in P$ and $\lim \delta(a, a^\mu) = 0$, whence for P the notion of limit in Definition 2.2 coincides with that already defined in P by means of δ. The details are omitted.

LEMMA 2.5. *Suppose $p, p^\mu, p_i, p_i^\mu \in P$ $(i = 1, \cdots, k; \mu = 1, 2, \cdots)$. If $p^\mu \leqq \sum [p_i^\mu; i = 1, \cdots, k]$, $\lim_\mu p_i^\mu = p_i$, $(p_1, \cdots, p_k)\perp$, $\lim p^\mu = p$, then $p \leqq \sum [p_i; i = 1, \cdots, k]$.*

Proof. Clearly $\delta(p^\mu, \sum_i p_i^\mu) = 0$, since $p^\mu \leqq \sum_i p_i^\mu$. But $\lim_\mu p_i^\mu = p_i$, $\lim p^\mu = p$, whence it follows from Lemma 2.4 that $\delta(p, \sum p_i) = 0$. Hence $p \leqq \sum p_i$ by the corollary to Axiom IV.

THEOREM 2.2. *If* $\lim_\mu p_i^\mu = p_i$ *for* $i = 1, \cdots, k$, *and if* $(p_1^\mu, \cdots, p_k^\mu)\perp$ *is false for* $\mu = 1, 2, \cdots$, *then* $(p_1, \cdots, p_k)\perp$ *is false.*

Proof. We establish first an auxiliary proposition: *There exist a subsequence* $(\mu_\nu; \nu = 1, 2, \cdots)$ *of the integers, an integer* $l < k$ *and distinct integers* $k_1, \cdots, k_l \in [1, \cdots, k]$ *such that*

$$(p_j^{\mu_\nu}; j = k_1, \cdots, k_l)\perp \qquad (\nu = 1, 2, \cdots),$$

and such that for each $m \neq k_1, \cdots, k_l$,

$$p_m^{\mu_\nu} \leqq \sum [p_j^{\mu_\nu}; j = k_1, \cdots, k_l] \qquad (\nu = 1, 2, \cdots).$$

Suppose the statement to be false. Then there exists $\nu = 1, 2, \cdots$ such that for $\mu > \nu$ and for $k_1, \cdots, k_l \in [1, \cdots, k]$ $(l < k)$ with $(p_{k_1}^\mu, \cdots, p_{k_l}^\mu)\perp$ it is true that there exists $m = 1, \cdots, k$ with $m \neq k_1, \cdots, k_l$, such that

$$(p_m^\mu, p_{k_1}^\mu, \cdots, p_{k_l}^\mu)\perp.$$

Let $\mu > \nu$ be fixed. Since $(p_1^\mu)\perp$, we may apply this result to $l = 1$, $k_1 = 1$, and obtain $m_1 = 1, \cdots, k$, $m_1 \neq 1$, such that $(p_{m_1}^\mu, p_1^\mu)\perp$. Another application yields $m_2 = 1, \cdots, k$, $m_2 \neq 1, m_1$, such that $(p_{m_2}^\mu, p_{m_1}^\mu, p_1^\mu)\perp$. Successive applications yield $(p_1^\mu, \cdots, p_k^\mu)\perp$, contrary to the hypothesis. Thus our proposition is established. Since $l \neq k$, there exists $m = 1, \cdots, k$ with $m \neq k_1, \cdots, k_l$. Then

$$p_m^{\mu_\nu} \leqq \sum (p_j^{\mu_\nu}; j = k_1, \cdots, k_l) \qquad (\nu = 1, 2, \cdots).$$

Let us suppose the conclusion of the theorem to be false, that is, $(p_1, \cdots, p_k)\perp$; then $(p_{k_1}, \cdots, p_{k_l})\perp$. Since $\lim_\nu p_i^{\mu_\nu} = p_i$ for $i = 1, \cdots, k$, we have, by Lemma 2.5,

$$p_m \leqq p_{k_1} + \cdots + p_{k_l},$$

contrary to $(p_1, \cdots, p_k)\perp$. This completes the proof.

COROLLARY. *If* $\lim_\mu p_i^\mu = p_i$ *for* $i = 1, \cdots, k$, *and if* $(p_1, \cdots, p_k)\perp$, *then there exists* $m = 1, 2, \cdots$ *such that* $(p_1^\mu, \cdots, p_k^\mu)\perp$ *for* $\mu > m$.

Proof. This is immediate from Theorem 2.2.

THEOREM 2.3. *Let* $a, a^\mu \in L$ *for* $\mu = 1, 2, \cdots$. *Suppose that* $d(a^\mu) = d(a) = k \geqq 1$ *for* $\mu = 1, 2, \cdots$, *and that* $(p_1, \cdots, p_k)$ *is a basis of* a. *Then*

(1) $$\lim a^\mu = a$$

is equivalent to

(2) *for each* $i = 1, \cdots, k$ and $\mu = 1, 2, \cdots$ *there exists* $p_i^\mu \leqq a^\mu$ *such that* $\lim_\mu p_i^\mu = p_i$ *for* $i = 1, \cdots, k$.

Proof. It is obvious that (1) implies (2). Suppose that (2) holds. Then by the corollary to Theorem 2.2, there exists $m = 1, 2, \cdots$ such that $(p_1^\mu, \cdots, p_k^\mu)\perp$ for $\mu > m$. Hence $p_1^\mu + \cdots + p_k^\mu = a^\mu$ $(\mu > m)$. Suppose $p = \lim p^\mu$, $p^\mu \leqq a^\mu$ $(\mu > n)$. Then $p^\mu \leqq p_1^\mu + \cdots + p_k^\mu$ $(\mu > m, n)$, and by Lemma 2.5 $p \leqq p_1 + \cdots + p_k = a$. Hence $L(a^\mu) \subset P_a$. Now suppose $p \leqq a$. Since $p \leqq p_1 + \cdots + p_k$, $(p_i)\perp$, we have

$$\lim_\mu \delta(p, p_1^\mu + \cdots + p_k^\mu) = 0$$

by Lemma 2.2. Thus by Lemma 2.1 there exists $(p^\mu; \mu = 1, 2, \cdots)$ with $p^\mu \leqq a^\mu$ such that $\lim p^\mu = p$. Hence $L(a^\mu) \supset P_a$. Therefore $L(a^\mu) = P_a$. But $\lim d(a^\mu) = d(a)$ by the hypothesis, whence (a^μ) is convergent, and $\lim a^\mu = a$ by Definition 2.2.

Remark. We have established in Theorem 2.3 a simple criterion for the convergence of an equi-dimensional sequence (a^μ); in geometrical terms this criterion states that (a^μ) converges to a when k sequences of basic points of a^μ converge to appropriate basic points of a.

It is now shown that L is a convergence-space.

THEOREM 2.4.

(a) *Let (a^μ) be convergent; then if (a^{μ_ν}) is a subsequence, (a^{μ_ν}) converges, and*

$$\lim_\nu a^{\mu_\nu} = \lim a^\mu.$$

(b) *The sequence $a^\mu = a$ is convergent, and* $\lim a^\mu = a$.

Proof of (a). Let $\lim a^\mu = a$. Since (a^μ) converges, there exists $m = 1, 2, \cdots$ such that $d(a^\mu) = d(a) \equiv k$ $(\mu > m)$. If $a = 0$, the desired result is evident. Suppose $a \neq 0$, and let $(p_1, \cdots, p_k)$ be a basis of a. By Theorem 2.3 there exist sequences $(p_i^\mu; \mu > m)$ such that $p_i^\mu \leqq a^\mu$ and $\lim_\mu p_i^\mu = p_i$ for $i = 1, \cdots, k$. Then $\lim_\nu p_i^{\mu_\nu} = p_i$ $(\mu_\nu > m)$, and by Theorem 2.3 $\lim_\nu a^{\mu_\nu} = a$ $(\mu_\nu > m)$. Hence $\lim_\nu a^{\mu_\nu} = a$ by the corollary to Definition 2.1.

Proof of (b). Let $p^\mu \leqq a$, $\lim p^\mu = p$. Then $p \leqq a$, since P_a is closed by Axiom IV. If $p \leqq a$, define $p^\mu \equiv p$, whence $p^\mu \leqq a^\mu$, and $\lim p^\mu = p$. Hence by the corollary to Theorem 2.1 $a = \lim a^\mu$.

THEOREM 2.5. *If $p^\mu \in P$, $a^\mu \in L$ for $\mu = 1, 2, \cdots$, and if* $\lim a^\mu = a$, $\lim p^\mu = p$, *then* $\lim \delta(p^\mu, a^\mu) = \delta(p, a)$.

Proof. This is an immediate consequence of Lemma 2.4 in view of the corollary to Theorem 2.2.

3. **Closed and open sets; neighborhoods.** A complete system of neighborhoods for L will be defined and studied. It is proved in Theorem 3.1 that L is a Hausdorff space.

By a *closed* set in L is meant a set that contains with every convergent sequence its limit. An *open* set is the complement in L of a closed set. The *closure* $\bar{S}$

of $S \subset L$ is the smallest closed subset of L containing S; its existence and uniqueness are easily proved. It is obvious that the intersection of any class of closed sets is closed, that the sum of any class of open sets is therefore open, that the sum of two closed sets is closed, and that the intersection of two open sets is therefore open. Indeed, it is clear that L is a *topological space* in the sense of Alexandroff and Hopf,[13] in the light of Theorem 2.4.

DEFINITION 3.1. Suppose $a \neq 0$, and let $(p_1, \cdots, p_k)$ be a basis of a. For each (real) $\epsilon > 0$, we define

$$U(a; \epsilon; p_1, \cdots, p_k) \equiv [b; d(b) = k, \delta(p_i, b) < \epsilon \ (i = 1, \cdots, k)].$$

For a fixed a the sets $U(a; \epsilon; p_1, \cdots, p_k)$ are called neighborhoods of a and are denoted generically by $U(a)$. The only neighborhood $U(0)$ of 0 is $[0]$.

LEMMA 3.1. *Let $a, a^\mu \epsilon L$ for $\mu = 1, 2, \cdots$. If $a \neq 0$, the following statements are equivalent:*

(a) *(a^μ) is convergent, and* $\lim a^\mu = a$;

(b) *for each $U(a)$ there exists $m = 1, 2, \cdots$ such that $a^\mu \epsilon U(a)$ for $\mu > m$;*

(c) *there exists a basis $(p_1, \cdots, p_k)$ of a such that for every $\epsilon > 0$ there exists $m = 1, 2, \cdots$ for which $a^\mu \epsilon U(a; \epsilon; p_1, \cdots, p_k)$ when $\mu > m$.*

If $a = 0$, then (a) *and* (b) *are equivalent.*

Proof. (a) $\rightarrow$ (b). For, let $U(a) = U(a; \epsilon; p_1, \cdots, p_k)$ be given. Since $\lim a^\mu = a$, there exist $n = 1, 2, \cdots$ and sequences $p_i^\mu \leqq a^\mu$, $\mu > n$, such that $\lim_{\mu > n} p_i^\mu = p_i$ for $i = 1, \cdots, k$. Hence there exists $m_1 = 1, 2, \cdots$ such that $\delta(p_i, p_i^\mu) < \epsilon$ for $\mu > m_1, n$ and $i = 1, \cdots, k$. Moreover, since (a^μ) is convergent, there exists $m_2 = 1, 2, \cdots$ such that $d(a^\mu) = k$ for $\mu > m_2, n$. Define $m \equiv \max(m_1, m_2, n)$, and suppose $\mu > m$. Then $d(a^\mu) = k$, and $\delta(p_i, a^\mu) \leqq \delta(p_i, p_i^\mu) < \epsilon$, whence $a^\mu \epsilon U(a)$.

Clearly (b) $\rightarrow$ (c). To prove that (c) $\rightarrow$ (a), let $\epsilon > 0$. Then there exists $m = 1, 2, \cdots$ such that $d(a^\mu) = k = d(a) \geqq 1$ and $\delta(p_i, a^\mu) < \epsilon$ for $\mu > m$. Thus $\lim_{\mu > m} \delta(p_i, a^\mu) = 0$ for $i = 1, \cdots, k$, whence by Lemma 2.1 there exist sequences $p_i^\mu \leqq a^\mu$, $\mu > m$, such that $\lim_{\mu > m} p_i^\mu = p_i$ for $i = 1, \cdots, k$. Hence $\lim a^\mu = a$ by Theorem 2.3 and the corollary to Definition 2.1.

The equivalence of (a) and (b) when $a = 0$ is obvious.

LEMMA 3.2. *Suppose $S \subset L$, $a \epsilon L$. If $a \neq 0$, then the following statements are equivalent:*

(a) *there exists a sequence $a^\mu \epsilon S$ $(\mu = 1, 2, \cdots)$ such that (a^μ) converges and* $\lim a^\mu = a$;

(b) *for every $U(a)$ there exists $b \epsilon S$ such that $b \epsilon U(a)$;*

(c) *there exists a basis $(p_1, \cdots, p_k)$ of a such that for every $\epsilon > 0$ there exists $b \epsilon S$ with $b \epsilon U(a; \epsilon; p_1, \cdots, p_k)$.*

If $a = 0$, then (a) *and* (b) *are equivalent.*

[13] Alexandroff and Hopf, loc. cit.

Proof. (a) $\rightarrow$ (b). This is obvious by Lemma 3.1 ((a) $\rightarrow$ (b)). Evidently (b) $\rightarrow$ (c). To prove that (c) $\rightarrow$ (a), let $(p_1, \cdots, p_k)$ be a basis of a, and let $\mu = 1, 2, \cdots$ be given. Then there exists $a^\mu \in S$ such that $a^\mu \in U(a; \mu^{-1}; p_1, \cdots, p_k)$. Hence $d(a^\mu) = k = d(a) \geqq 1, \delta(p_i, a^\mu) < \mu^{-1}$. By Lemma 2.1 there exist $p_i^\mu \leqq a^\mu$ $(\mu = 1, 2, \cdots; i = 1, \cdots, k)$, such that $\lim_\mu p_i^\mu = p_i$. Thus by Theorem 2.3, $\lim a^\mu = a$. If $a = 0$, (a) and (b) are obviously equivalent.

Lemma 3.3. *Let $G \subset L$ be open, suppose $a \in G$, $a \neq 0$, and let $(p_1, \cdots, p_k)$ be a basis of a. Then there exists $\epsilon > 0$ with $U(a; \epsilon; p_1, \cdots, p_k) \subset G$.*

Proof. Since $a \epsilon' L - G$, and since $L - G$ is closed, it is false by Lemma 3.2 that for each $\epsilon > 0$ there exists $b \in L - G$ with $b \in U(a; \epsilon; p_1, \cdots, p_k)$. Hence there exists $\epsilon > 0$ such that every b in $L - G$ is not in $U(a; \epsilon; p_1, \cdots, p_k)$. Thus $U(a; \epsilon; p_1, \cdots, p_k) \cdot (L - G) = \Theta$, whence $U(a; \epsilon; p_1, \cdots, p_k) \subset G$.

Corollary. *Every open set $G \subset L$ contains with each of its elements a a neighborhood $U(a)$.*

The proof is obvious.

Lemma 3.4. *If $a \in L$, then each neighborhood $U(a)$ is open.*

Proof. If $a = 0$, then $U(a) = [0]$, whence it is easily seen that $L - U(a)$ is closed; thus $U(a)$ is open. Suppose that $a \neq 0$. Let $U(a) = U(a; \epsilon; p_1, \cdots, p_k)$ be given; that is, suppose $\epsilon > 0$ and let $(p_1, \cdots, p_k)$ be a basis of a. We shall prove that $L - U(a)$ is closed. Suppose that $b^\mu \in L - U(a)$, $\lim b^\mu = b$. It is to be shown that $b \in L - U(a)$. Suppose this to be false, i.e., $b \in U(a)$. Then $d(b) = d(a) = k \geqq 1$, and $\delta(p_i, b) < \epsilon$ for $i = 1, \cdots, k$. Let ϵ_1 be so chosen that $\delta(p_i, b) < \epsilon_1 < \epsilon$, and define $\epsilon_2 \equiv \epsilon - \epsilon_1 > 0$. By Theorem 2.5 and Definition 2.2 there exists $m = 1, 2, \cdots$ such that for $\mu > m$, $d(b^\mu) = d(b) = k \geqq 1$, and

$$| \delta(p_i, b^\mu) - \delta(p_i, b) | < \epsilon_2,$$

whence

$$| \delta(p_i, b^\mu) | \leqq | \delta(p_i, b^\mu) - \delta(p_i, b) | + | \delta(p_i, b) | < \epsilon_2 + \epsilon_1 = \epsilon.$$

Thus $b^\mu \in U(a)$ for $\mu > m$, contrary to $b^\mu \in L - U(a)$. This contradiction shows that $b \in L - U(a)$, and the proof is complete.

Lemma 3.5. *If $b \in U(a)$, then there exists $U(b)$ such that $U(b) \subset U(a)$.*

Proof. This is obvious by the corollary to Lemma 3.3 since $U(a)$ is open by Lemma 3.4.

Lemma 3.6. *Suppose $a, b \in L$, $a \neq b$. Then there exist $U(a)$, $U(b)$ such that $U(a) \cdot U(b) = \Theta$.*

Proof. If $a = 0$ or $b = 0$, the conclusion is trivial. Suppose $a, b \neq 0$, and assume the conclusion to be false. Let $(\epsilon^\mu; \mu = 1, 2, \cdots)$ be a sequence of positive real numbers such that $\lim \epsilon^\mu = 0$, and let $(p_1, \cdots, p_k)$ and $(q_1, \cdots, q_l)$

be bases of a and b respectively. Then there exists a sequence $(c^\mu; \mu = 1, 2, \cdots)$ such that

$$c^\mu \in U(a; \epsilon^\mu; p_1, \cdots, p_k) \cdot U(b; \epsilon^\mu; q_1, \cdots, q_l) \qquad (\mu = 1, 2, \cdots).$$

Thus $d(a) = d(c^\mu) = d(b) \geqq 1$, and $\delta(p_i, c^\mu)$, $\delta(q_i, c^\mu) < \epsilon^\mu$ for $\mu = 1, 2, \cdots$. Therefore

$$\lim_\mu \delta(p_i, c^\mu) = \lim_\mu \delta(q_i, c^\mu) = 0,$$

and $a = \lim c^\mu = b$ by Lemma 2.1, Theorem 2.3. This contradiction completes the proof.

THEOREM 3.1. *The lattice L is a Hausdorff space.*[14]

Proof. The neighborhood axioms as applied to the sets $U(a)$, $a \in L$, follow immediately from Lemmas 3.4, 3.5, the corollary to Lemma 3.3, and Lemma 3.6. The Hausdorff separation axiom is Lemma 3.6.

DEFINITION 3.2. For each $k = 0, \cdots, N$ define $L_k \equiv [a; d(a) = k]$.

THEOREM 3.2. *Every set L_k is closed and open.*

Proof. From Definition 2.2 it is clear that each L_k is closed. Since $L_k = L - \sum [L_i; i \neq k]$, L_k is also open.

THEOREM 3.3. *Let for each $a \neq 0$ a basis $(p_i(a); i = 1, \cdots, k)$, $k = d(a)$, be given. The sets $U(a; \epsilon; p_1(a), \cdots, p_k(a))$, $a \in L$, $\epsilon > 0$, together with $U(0) = [0]$ form a topological basis*[15] *of L.*

Proof. Let $G \subset L$ be open. Since

$$G = G \cdot L = G \cdot (L_0 + \cdots + L_N) = G \cdot L_0 + \cdots + G \cdot L_N,$$

and since each $G \cdot L_i$ is open, it suffices to prove that each $G \cdot L_i$ is the sum of neighborhoods of the desired form. Suppose $i \geqq 1$ and $a \in G \cdot L_i$; by Lemma 3.3 there exists $\epsilon(a) > 0$ such that

$$U(a) \equiv U(a; \epsilon(a); p_1(a), \cdots, p_k(a)) \subset G \cdot L_i \qquad (i = 1, \cdots, N).$$

Of course $U(0) = [0] \subset G \cdot L_0$ if $0 \in G$. Thus

$$G \cdot L_i \subset \sum [U(a); a \in G \cdot L_i] \subset G \cdot L_i \qquad (i = 0, \cdots, N),$$

and the proof is complete.

4. **Continuity in L.** A brief investigation is given here of the continuity of the operation $a + b$, the function $d(a)$, and the relation "$(a, b) \perp$ is false".

THEOREM 4.1. *If $a^\mu \in L$ $(\mu = 1, 2, \cdots)$, and $a = \sum L(a^\mu)$, then*

$$\liminf_{\mu \to \infty} d(a^\mu) \geqq d(a).$$

[14] Alexandroff and Hopf, op. cit., p. 67.

[15] A topological basis is a family of open sets such that every open set is the sum of sets of the family.

Proof. If $a = 0$, the result is trivial. Suppose $a \neq 0$, and let $(p_1, \cdots, p_k)$, $k = d(a)$, be a basis of a. Then by Theorem 2.1 there exist $n = 1, 2, \cdots$ and sequences $p_i^\mu \leqq a^\mu$ ($\mu > n$; $i = 1, \cdots, k$), such that $\lim_\mu p_i^\mu = p_i$. Moreover, by the corollary to Theorem 2.2, there exists $m = 1, 2, \cdots$ such that $(p_1^\mu, \cdots, p_k^\mu)\perp$ for $\mu > m, n$. Hence by Lemma 1.1

$$d(a^\mu) \geqq d(\sum_i p_i^\mu) = \sum_i d(p_i^\mu) = k \qquad (\mu > m, n),$$

whence the conclusion follows.

THEOREM 4.2. *If* $(a_1^\mu, \cdots, a_n^\mu)\perp$ *is false for* $\mu = 1, 2, \cdots$, *and if* $\lim_\mu a_i^\mu = a_i$, *then* $(a_1, \cdots, a_n)\perp$ *is false.*

Proof. Suppose that $(a_1, \cdots, a_n)\perp$, whence $a_i \neq 0$, and let $(p_{ij}; j = 1, \cdots, k_i)$, $k_i = d(a_i)$, be a basis of a_i. Now there exist $l = 1, 2, \cdots$ and sequences $p_{ij}^\mu \leqq a_i^\mu$ ($\mu > l; j = 1, \cdots, k_i; i = 1, \cdots, n$), such that $\lim_\mu p_{ij}^\mu = p_{ij}$. Moreover, by the corollary to Theorem 2.2, since $(p_{ij}, j = 1, \cdots, k_i; i = 1, \cdots, n)\perp$, there exists $m = 1, 2, \cdots$ such that $(p_{ij}^\mu; j = 1, \cdots, k_i; i = 1, \cdots, n)\perp$ and $d(a_i^\mu) = d(a_i)$ for $\mu > m, l$. Hence $a_i^\mu = \sum [p_{ij}^\mu; j = 1, \cdots, k_i]$, and $(a_i^\mu; i = 1, \cdots, n)\perp$ for $\mu > m, l$. This contradiction completes the proof.

COROLLARY. *If* $(a_1, \cdots, a_n)\perp$, *and if* $\lim_\mu a_i^\mu = a_i$ *for* $i = 1, \cdots, n$, *then there exists* $m = 1, 2, \cdots$ *such that* $(a_1^\mu, \cdots, a_n^\mu)\perp$ *for* $\mu > m$.

Proof. This is immediate from Theorem 4.2.

Remark. The results just proved extend those contained in Theorem 2.2 and its corollary from P to L.

THEOREM 4.3. *Let* $\lim a^\mu = a$, *and* $\lim b^\mu = b$. *Then*

$$\lim (a^\mu + b^\mu) = a + b$$

if and only if

$$\lim d(a^\mu + b^\mu) = d(a + b).$$

Proof. The case $a = 0$ or $b = 0$ is trivial. Suppose $a, b \neq 0$. The forward implication is obvious. To prove the converse, let $(p_1, \cdots, p_k)$ and $(q_1, \cdots, q_l)$ be bases of a and b respectively. Then there exist $n = 1, 2, \cdots$ and sequences $p_i^\mu \leqq a^\mu$, $q_j^\mu \leqq b^\mu$ ($\mu > n; i = 1, \cdots, k; j = 1, \cdots, l$) such that $\lim_\mu p_i^\mu = p_i$, $\lim_\mu q_j^\mu = q_j$. Application of Theorems 1.3 and 2.3 yields the desired result.

COROLLARY. *If* $\lim a^\mu = a$, $\lim b^\mu = b$, *and if* $(a, b)\perp$, *then*

$$\lim (a^\mu + b^\mu) = a + b.$$

Proof. By the corollary to Theorem 4.2 there exists $m = 1, 2, \cdots$ such that $(a^\mu, b^\mu)\perp$ for $\mu > m$. Hence by Lemma 1.1

$$d(a^\mu + b^\mu) = d(a^\mu) + d(b^\mu),$$

and

$$\lim d(a^\mu + b^\mu) = d(a) + d(b) = d(a + b).$$

The conclusion follows from Theorem 4.3.

We close with a brief survey of questions suggested by the theory which we have developed. Weakening of the axiom that P be a metric space might be considered. When P is metric, as in this paper, it may be possible to prove stronger properties for L than that it is a mere Hausdorff space; the most important problem here is to find conditions under which L is a metric space, as is the case with projective and affine spaces. Implications for L of further topological assumptions on P, for example, connectedness, completeness, compactness, etc., remain to be investigated. Finally, it might be of interest to ascertain how much continuity is possessed by the relations $(a, b)M$, $a \parallel b$, and the operation $a \cdot b$, under our Axioms I–IV.

ILLINOIS INSTITUTE OF TECHNOLOGY.

THE ARITHMETICAL THEORY OF BIRKHOFF LATTICES

By R. P. Dilworth

1. **Introduction and summary.** In the development of lattice theory considerable work has been devoted to the study of the arithmetical properties of modular and distributive lattices. Indeed most of the decomposition theorems of abstract algebra have been extended to these more general domains. Nevertheless, there are lattices with very simple arithmetical properties which come under neither of these classifications. For example, the lattices with unique irreducible decompositions, which were studied by the author in a previous paper [3][1] satisfy the Birkhoff condition[2] which is even less restrictive than the modular axiom. Furthermore, there are important algebraic systems which give rise to non-modular, Birkhoff lattices. Thus, since every exchange lattice (Mac Lane [4]) is a Birkhoff lattice, the systems which satisfy Mac Lane's exchange axiom form lattices of the type in question. In this paper we shall study the arithmetical structure of general Birkhoff lattices and in particular determine necessary and sufficient conditions that certain important arithmetical properties hold.

In §§2–4 we characterize the irreducible decompositions in terms of the structure of the lattice and apply the results to determine necessary and sufficient conditions that the number of irreducible components be unique for each element of the lattice. The main result is the following:

Let $\mathfrak{S}$ be a Birkhoff lattice in which every quotient lattice is Archimedean. Then the number of irreducible components is unique for each element a of $\mathfrak{S}$ if and only if the sublattice generated by the elements covering a is a dense, modular sublattice of $\mathfrak{S}$.

§5 contains some methods for constructing Birkhoff lattices in which given arithmetical conditions hold. In §6 we treat the problem of determining the conditions that a set of irreducible components of an element must satisfy in order that it may be extended into a reduced representation. This problem is given a particularly simple solution in the case of a Birkhoff lattice in which the number of components is unique.

2. **Notation and definitions.** Throughout the paper $\mathfrak{S}$ will denote a lattice of elements $a, b, c, \cdots$ and unit element u in which every quotient lattice is Archimedean (Ore [5]). Union, cross-cut, and lattice division will be denoted by $(\ , \ )$, $[\ , \ ]$, and $\supset$ respectively. a is said to *cover* b (in symbols $a > b$) if $a \supset b$, $a \neq b$ and $a \supset x \supset b$ implies $a = x$ or $x = b$. Elements covered by

Received October 9, 1940. The author is a Sterling Research Fellow at Yale University.

[1] Numbers in brackets refer to the references at the end of the paper.

[2] See Definition 2.1.

the unit element u are said to be *simple*. If $\mathfrak{S}$ has a null element z, the elements which cover z will be called *points*. $\mathfrak{S}$ is said to be *atomic* if every element is a cross-cut of simple elements. Similarly, if every element is a union of points, $\mathfrak{S}$ is said to be a *point* lattice.

An element x is said to be *cross-cut irreducible* (or simply *irreducible*) if $x = [x_1, x_2]$ implies $x = x_1$ or $x = x_2$. x is said to be *union irreducible* if $x = (x_1, x_2)$ implies $x = x_1$ or $x = x_2$. A set of elements $x_1, \cdots, x_n$ is said to be *cross-cut independent* if $x_i \not\supset [x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n]$ $(i = 1, \cdots, n)$. Similarly, $x_1, \cdots, x_n$ are *union independent* if $(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n) \not\supset x_i$ $(i = 1, \cdots, n)$.

Since the ascending chain condition holds in $\mathfrak{S}$, each element a of $\mathfrak{S}$ may be expressed as a cross-cut of irreducibles $a = [q_1, \cdots, q_k]$. If $q_1, \cdots, q_k$ are cross-cut independent, the representation is said to be *reduced*. Clearly any representation can be reduced by dropping out suitable members. If an irreducible q occurs in a reduced decomposition of a, q is called a *component* of a.

Let $\mathfrak{A}_1, \cdots, \mathfrak{A}_n$ be sublattices of $\mathfrak{S}$ with common null element a. The sublattice generated by $\mathfrak{A}_1, \cdots, \mathfrak{A}_n$ is the *direct sum* of $\mathfrak{A}_1, \cdots, \mathfrak{A}_n$ ($\mathfrak{A} = \mathfrak{A}_1 + \mathfrak{A}_2 + \cdots + \mathfrak{A}_n$) if $x \in \mathfrak{A}$ implies $x = (x_1, \cdots, x_n)$, where $x_i \in \mathfrak{A}_i$, $(x, y) = ((x_1, y_1), \cdots, (x_n, y_n))$, $[x, y] = ([x_1, y_1], \cdots, [x_n, y_n])$ and $x \supset y$ if and only if $x_i \supset y_i$ $(i = 1, \cdots, n)$. If $\mathfrak{A}_2$ is the quotient lattice p/a where $p > a$, we write $\mathfrak{A} = \mathfrak{A}_1 + \mathfrak{A}_2 = \mathfrak{A}_1 + p$ and say that $\mathfrak{A}_1 + p$ is the direct sum of $\mathfrak{A}_1$ and the point p.

Definition 2.1. A lattice $\mathfrak{S}$ is said to be a *Birkhoff*[3] lattice of type I (or simply *Birkhoff* lattice) if $a > [a, b]$ implies $(a, b) > b$. $\mathfrak{S}$ is said to be a *Birkhoff* lattice of type II if $(a, b) > b$ implies $a > [a, b]$.

If $\mathfrak{S}$ is a Birkhoff lattice of type I, a *rank* function ρa may be defined over $\mathfrak{S}$ with the properties (i) $\rho u = 0$, (ii) $\rho a = \rho b + 1$ if $b > a$, (iii) $\rho(a, b) + \rho[a, b] \geqq \rho a + \rho b$. If $\mathfrak{S}$ is a Birkhoff lattice of type II, then in place of (iii) ρ satisfies (iii)′ $\rho(a, b) + \rho[a, b] \leqq \rho a + \rho b$.

The following lemma proved by G. Birkhoff in [1] relates the Birkhoff conditions to modularity.

Lemma 2.1. *$\mathfrak{S}$ is modular if and only if it is a Birkhoff lattice of type* I *and type* II.

3. **Structure characterization of reduced decompositions.** We begin by proving a basic lemma from which most of the properties of Birkhoff lattices may be derived.

Fundamental Lemma. *Let $\mathfrak{S}$ be a Birkhoff lattice. Let $\mathfrak{A}$ be a complete*[4] *modular sublattice of $\mathfrak{S}$ with unit element v and null element a. Then if $p > a$ and $v \not\supset p$, the sublattice generated by $\mathfrak{A}$ and p is the direct sum $\mathfrak{A} + p$. $\mathfrak{A} + p$ is also complete in $\mathfrak{S}$.*

[3] G. Birkhoff (*Lattice Theory*, Amer. Math. Soc. Colloquium Publications, vol. 25, New York, 1940) uses the terms *upper* and *lower semi-modular* lattices for Birkhoff lattices of types I and II. The original terminology of Klein has been used here.

[4] A sublattice $\mathfrak{A}$ of $\mathfrak{S}$ is said to be complete in $\mathfrak{S}$ if $a > b$ in $\mathfrak{A}$ implies $a > b$ in $\mathfrak{S}$.

Proof. Let a_1 and a_2 be elements of $\mathfrak{A}$. If $(a_1, p) = (a_2, p)$, then $(a_1, p) \supset (a_1, a_2) \supset a_1$. Since $(a_1, p) > a_1$, either $(a_1, p) = (a_1, a_2)$ or $a_1 \supset a_2$. But if $(a_1, p) = (a_1, a_2)$, then $v \supset p$, and this is contrary to assumption. Hence $a_1 \supset a_2$ and similarly $a_2 \supset a_1$, $a_1 = a_2$. Thus $(a_1, x_1) = (a_2, x_2)$, $a_1, a_2 \in \mathfrak{A}$, $x_1, x_2 \in p/a$ implies $a_1 = a_2$, $x_1 = x_2$.

Clearly $((a_1, x_1), (a_2, x_2)) = ((a_1, a_2), (x_1, x_2))$.

Now $\rho[(a_1, p), (a_2, p)] \geqq \rho(a_1, p) + \rho(a_2, p) - \rho(a_1, a_2, p) = \rho a_1 + \rho a_2 - 2 - \rho(a_1, a_2) + 1 = \rho[a_1, a_2] - 1 = \rho([a_1, a_2], p)$ since $\mathfrak{A}$ is modular. Hence $\rho[(a_1, p), (a_2, p)] \geqq \rho([a_1, a_2], p)$. Since $[(a_1, p), (a_2, p)] \supset ([a_1, a_2], p)$, it follows that $[(a_1, p), (a_2, p)] = ([a_1, a_2], p)$. Also $([a_1, a_2], p) \supset [(a_1, p), a_2] \supset [a_1, a_2]$ and $a_2 \not\supset p$ imply $[(a_1, p), a_2] = [a_1, a_2]$. Hence $[(a_1, x_1), (a_2, x_2)] = ([a_1, a_2], [x_1, x_2])$, and the lemma is proved.

By repeated application of the Fundamental Lemma we have the following lemmas:

Lemma 3.1. *Let $\mathfrak{S}$ be a Birkhoff lattice and let $a_1, \cdots, a_n > a$. Then each union independent set of the a_i is contained in a maximal union independent set.*

Lemma 3.2. *Let $\mathfrak{S}$ be a Birkhoff lattice and let $a_1, \cdots, a_n > a$. Then every union independent set of the a_i generates a Boolean algebra.*

Lemma 3.3. *Let $\mathfrak{S}$ be a Birkhoff lattice and let $a_1, \cdots, a_n > a$. Then any two maximal independent sets of the a_i have the same number of elements and any element of one set may be replaced by a suitably chosen element of the other.*

For if $a_1, \cdots, a_k$ and $a_1', \cdots, a_l'$ are the two maximal independent sets of the a_i, we may replace a_i by an a_j' such that $(a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_n) \not\supset a_j'$.

Lemma 3.4. *Let $\mathfrak{S}$ be a Birkhoff lattice and let $a > s_1, \cdots, s_k$ where $s_1, \cdots, s_k$ are cross-cut independent and $\rho[s_1, \cdots, s_k] - \rho(a) = k$. Then $s_1, \cdots, s_k$ generate a Boolean algebra of length k.*

For if $p_i = [s_1, \cdots, s_{i-1}, s_{i+1}, \cdots, s_k]$, then $p_i > [s_1, \cdots, s_k]$ and $p_1, \cdots, p_k$ are independent. Thus $s_i = (p_1, \cdots, p_{i-1}, p_{i+1}, \cdots, p_k)$ and hence the sublattice generated by $s_1, \cdots, s_k$ is simply the Boolean algebra generated by $p_1, \cdots, p_k$.

If a is an element of $\mathfrak{S}$, let u_a denote the union of the elements which cover a. We associate with each element a the quotient lattice $\mathfrak{S}_a$ of all elements x such that $u_a \supset x \supset a$.

Definition 3.1. An element $c \neq u_a$ of $\mathfrak{S}_a$ is said to be *characteristic* if there exists an irreducible q such that $q \supset c$ and q divides exactly the same points of $\mathfrak{S}_a$ as c.

Each irreducible component of a divides at least one characteristic element c. For the union of the points of $\mathfrak{S}_a$ divisible by q clearly has the properties of Definition 3.1.

We show now that each reduced decomposition of a into irreducibles gives at least one representation of a as a reduced cross-cut of characteristic elements of $\mathfrak{S}_a$ and conversely.

THEOREM 3.1. *An element a of $\mathfrak{S}_a$ has a reduced decomposition $a = [q_1, \cdots, q_k]$ where $q_1, \cdots, q_k$ are irreducible, if and only if a has a reduced representation $a = [c_1, \cdots, c_k]$, where $c_1, \cdots, c_k$ are characteristic elements of $\mathfrak{S}_a$ such that $q_i \supset c_i$.*

Proof. Let $a = [q_1, \cdots, q_k]$ be an irreducible decomposition of a and let $c_1, \cdots, c_k$ be a set of associated characteristic elements. Then $a = [q_1, \cdots, q_k] \supset [c_1, \cdots, c_k] \supset a$. Hence $a = [c_1, \cdots, c_k]$. Now since $q_1, \cdots, q_k$ are cross-cut independent, $[q_1, \cdots, q_{i-1}, q_{i+1}, \cdots, q_k] \supset p_i$, where $p_i > a$. But then $[c_1, \cdots, c_{i-1}, c_{i+1}, \cdots, c_k] \supset p_i$ by Definition 3.1. Hence $c_i \not\supset [c_1, \cdots, c_{i-1}, c_{i+1}, \cdots, c_k]$ and the representation $a = [c_1, \cdots, c_k]$ is reduced.

Conversely, let $a = [c_1, \cdots, c_k]$ be a reduced representation of a in terms of characteristic elements and $q_1, \cdots, q_k$ be a set of associated irreducibles. If $[q_1, \cdots, q_k] \neq a$, then $[q_1, \cdots, q_k] \supset p > a$. Since $q_i \supset p$ implies $c_i \supset p$, we have $a = [c_1, \cdots, c_k] \supset p > a$, and this is impossible. Hence $a = [q_1, \cdots, q_k]$. If $q_i \supset [q_1, \cdots, q_{i-1}, q_{i+1}, \cdots, q_k]$, then $a = [q_1, \cdots, q_{i-1}, q_{i+1}, \cdots, q_k] \supset [c_1, \cdots, c_{i-1}, c_{i+1}, \cdots, c_k] \supset a$. This contradicts the assumption that $c_1, \cdots, c_k$ are cross-cut independent.

It will be noted that Theorem 3.1 is independent of the Birkhoff condition.

In view of Theorem 3.1 the structure characterization of the irreducible decompositions will be accomplished if we determine the characteristic elements of $\mathfrak{S}_a$ in terms of the lattice structure. We need the following lemma:

LEMMA 3.5. *$\mathfrak{S}_a$ is a complemented, atomic lattice.*

Proof. Let $b \in \mathfrak{S}_a$ and let $p_1, \cdots, p_k$ be a maximal union independent set of elements covering a and divisible by b. Imbed $p_1, \cdots, p_k$ in a maximal union independent set of points $p_1, \cdots, p_n$ (Lemma 3.1). Let $b' = (p_{k+1}, \cdots, p_n)$. Then clearly $(b, b') = u_a$. Suppose that $[b, b'] \neq a$. Then $[b, b'] \supset p > a$. But since $b \supset p$, we have $(p_1, \cdots, p_k) \supset p$ and $a = [(p_1, \cdots, p_k), (p_{k+1}, \cdots, p_n)] \supset p$ (Lemma 3.2), and this is impossible. Hence $[b, b'] = a$ and thus $\mathfrak{S}_a$ is complemented. Let now b be an irreducible element of $\mathfrak{S}_a$. Then $(b, p_{k+1}), \cdots, (b, p_n) > b$ and hence $(b, p_{k+1}) = \cdots = (b, p_n)$. But then $u_a = (b', b) > b$, and thus the only irreducible elements of $\mathfrak{S}_a$ are the simple elements.

If $b \in \mathfrak{S}_a$, an arbitrary complement of b in $\mathfrak{S}_a$ will be denoted by b'.

THEOREM 3.2. *Let $\mathfrak{S}$ be a Birkhoff lattice. Then an element $c \in \mathfrak{S}_a$ is characteristic if and only if there exists a divisor x of c such that $(c', x) > x$ and $[c', x] = a$ for every c'.*

Proof. Let us first assume that such an element x exists. Then $(u_a, x) = ((c, c'), x) = (c', x)$. Let q be an irreducible such that $q \supset x$, $q \not\supset (u_a, x)$. Since $q \supset x \supset c$, q divides every point of $\mathfrak{S}_a$ which c divides. Now let $q \supset p$. Then $x \supset p$ since if $x \not\supset p$ we would have $(c', x) \supset (x, p) \supset x$ and $(x, p) \neq x$. Hence $(c', x) = (x, p)$ and $q \supset (x, p) = (c', x)$, and this contradicts the definition of q. Now if $c \not\supset p$, then $c' \supset p$ for some c'. But then $a = [c', x] \supset p$, and this is impossible. Hence $q \supset p$ implies $c \supset p$ and c is thus characteristic.

On the other hand, let c be characteristic and let q be an irreducible associated with c. Then $(q, c') > q$ for every c'. For there is a point p such that $c' \supset p$, $c \not\supset p$ since otherwise we would have $c' = a$ and $c = u_a$. But then $(q, c') = (q, c, c') = (q, u_a) = (q, p) > q$. Now if $[c', q] \neq a$, then $[c', q] \supset p > a$, and hence $c' \supset p$, $q \supset p$. But then $c \supset p$ and hence $a = [c, c'] \supset p$; this is impossible. Thus $[c', q] = a$ for every c'.

COROLLARY 3.2. *Each simple element of $\mathfrak{S}_a$ is characteristic.*

For we may take x to be the element c itself.

COROLLARY 3.3. *If k is the maximal number of union independent elements covering a, then a has a reduced decomposition into irreducibles with k components.*

For by Theorem 3.1 and Corollary 3.2 a has a reduced representation as a cross-cut of k characteristic elements of $\mathfrak{S}_a$.

If $\mathfrak{S}$ is modular, then $\mathfrak{S}$ is a Birkhoff lattice of type II by Lemma 2.1 and hence by Theorem 3.2 we must have $c' > a$. But then $u_a > c$, and every characteristic element of $\mathfrak{S}_a$ is simple in $\mathfrak{S}_a$. Hence by the dual of Lemma 3.3 the number of components in any two irreducible decompositions is the same and each component of one decomposition may be replaced by a suitably chosen component of the other. Thus for modular lattices in which every quotient lattice is Archimedean we see that the Kurosch-Ore decomposition theorem rests on a familiar exchange property of independent bases.

4. **Unicity of the number of components.** In order to investigate the structure of Birkhoff lattices in which the number of components is unique we prove first a lemma which sharpens Corollary 3.3.

LEMMA 4.1. *Let $\mathfrak{S}$ be a Birkhoff lattice. Then a reduced representation $a = [q_1, \cdots, q_k]$ has the maximal number of components if and only if the characteristic elements belonging to the q_i are the simple elements of the Boolean algebra generated by a maximal union independent set of points of $\mathfrak{S}_a$.*

Proof. Clearly each such set of characteristic elements gives a reduced representation of a by Lemma 3.2 and Theorem 3.1. Now if $a = [q_1, \cdots, q_k]$ is a reduced representation of a, let $q_i' = [q_1, \cdots, q_{i-1}, q_{i+1}, \cdots, q_k]$. Then since $q_i' \neq a$, we must have $q_i' \supset p_i$ where $p_i > a$. $p_1, \cdots, p_k$ are union independent since if $(p_1, \cdots, p_{i-1}, p_{i+1}, \cdots, p_k) \supset p_i$, then $q_i \supset (q_1', \cdots, q_{i-1}', q_{i+1}', \cdots, q_k') \supset p_i$. Hence $a = [q_i, q_i'] \supset p_i$, and this contradicts $p_i > a$. Let $b_i = (p_1, \cdots, p_{i-1}, p_{i+1}, \cdots, p_k)$. Then $q_i \supset b_i$. Since $p_1, \cdots, p_k$ are union independent, $k \leqq n$ where n is the maximal number of union independent points of $\mathfrak{S}_a$. But if the representation $a = [q_1, \cdots, q_k]$ has the maximal number of components, $k = n$ and $p_1, \cdots, p_k$ are a maximal union independent set of points of $\mathfrak{S}_a$ and hence generate a Boolean algebra with simple elements $b_1, \cdots, b_k$. $b_1, \cdots, b_k$ are clearly simple in $\mathfrak{S}_a$. Now let $c_1, \cdots, c_k$ be characteristic elements associated with $q_1, \cdots, q_k$. Then since $q_i \supset b_i$ we have

$u_a \supset c_i \supset b_i$ by Definition 3.1. But $u_a > b_i$ and $u_a \neq c_i$. Hence $c_i = b_i$ and the lemma is proved.

COROLLARY 4.1. *The maximal number of components in the reduced representations of a as a cross-cut of irreducibles is equal to the length of* $\mathfrak{S}_a$.

THEOREM 4.1. *Let* $\mathfrak{S}$ *be a Birkhoff lattice. Then the number of components in the reduced representations of a as a cross-cut of irreducibles is unique if and only if* $\mathfrak{S}_a$ *is modular and the characteristic elements are simple.*

Proof. If $\mathfrak{S}_a$ is modular and the only characteristic elements of $\mathfrak{S}_a$ are the simple elements, then since $\mathfrak{S}_a$ is a Birkhoff lattice of type II the number of elements in the representations of a as a cross-cut of characteristic elements is unique by the dual of Lemma 3.3. Hence the number of components in the irreducible decomposition of a is unique by Theorem 3.1.

On the other hand, if the number of components in the irreducible decompositions of a is unique, then each cross-cut independent set of simple elements of $\mathfrak{S}_a$ generates a Boolean algebra whose length is equal to the number of elements in the set. For if $u_a > s_1, \cdots, s_k$ and $s_1, \cdots, s_k$ do not generate a Boolean algebra, then $\rho[s_1, \cdots, s_k] - \rho u_a = l > k$ by Lemma 3.4. But then by adding $n - l$ simple elements $s_{l+1}, \cdots, s_n$ we have $a = [s_1, \cdots, s_k, s_{l+1}, \cdots, s_n]$, where n is the length of $\mathfrak{S}_a$. This, however, contradicts Lemma 4.1. Now let $b \in \mathfrak{S}_a$. Then $b = [s_1, \cdots, s_k]$ (Lemma 3.5), where $s_1, \cdots, s_k$ generate a Boolean algebra. If $s \not\supset b$, then $s, s_1, \cdots, s_k$ must generate a Boolean algebra of length $k + 1$ and hence $b > [s, b]$. Let $(x, y) > y$ where $x, y \in \mathfrak{S}_a$. Then by Lemma 3.5 there is a simple element s of $\mathfrak{S}_a$ such that $s \not\supset (x, y)$, $s \supset y$. Hence $y = [(x, y), s]$. Since $s \not\supset x$ we have $x > [x, s] = [x, (x, y), s] = [x, y]$. Thus $\mathfrak{S}_a$ is a Birkhoff lattice of type II and hence is modular by Lemma 2.1. The characteristic elements of $\mathfrak{S}_a$ are simple by Lemma 4.1.

To complete the proof of the theorem mentioned in the introduction we must prove first a lemma on modular sublattices of a lattice. Let $\mathfrak{A}$ and $\mathfrak{B}$ be quotient lattices of $\mathfrak{S}$ with unit elements a_1, b_1 and null elements a_2, b_2 respectively. Then if $a_1 \in \mathfrak{B}$ and $b_2 \in \mathfrak{A}$ we say that $\mathfrak{B}$ is an *extension* of $\mathfrak{A}$.

LEMMA 4.2. *Let* $\mathfrak{A}_1, \cdots, \mathfrak{A}_k$ *be quotient lattices of* $\mathfrak{S}$ *such that* $\mathfrak{A}_{i+1}$ *is an extension of* $\mathfrak{A}_i$. *Then the set sum* $\mathfrak{S}_k$ *of the lattice* $\mathfrak{A}_1, \cdots, \mathfrak{A}_k$ *is a sublattice of* $\mathfrak{S}$ *and is modular if and only if* $\mathfrak{A}_1, \cdots, \mathfrak{A}_k$ *are modular.*

Proof. The unit and null elements of $\mathfrak{A}_i$ will be denoted by a_i and b_i respectively. Let a and b be elements in the set sum $\mathfrak{S}_k$. Then $a \in \mathfrak{A}_i$ and $b \in \mathfrak{A}_j$ where we may assume $i \leqq j$. Then $(a, b) = (a, b, b_j)$ since $b \supset b_j$. But then $a_j \supset (a, b, b_j) \supset b_j$ and hence $(a, b) \in \mathfrak{A}_j$. Similarly $[a, b] \in \mathfrak{A}_i$. Hence $\mathfrak{S}_k$ is a sublattice of $\mathfrak{S}$.

Now let $\mathfrak{A}_1, \cdots, \mathfrak{A}_k$ be modular and suppose that it has been shown that $\mathfrak{S}_{k-1}$ is modular. We prove then that $\mathfrak{S}_k$ is modular and the lemma follows by induction. Let a, b, c be any three elements of $\mathfrak{S}_k$ such that $a \supset b$. We have four non-trivial cases to consider.

(1) $a, c \in \mathfrak{A}_k$, $b \in \mathfrak{S}_{k-1}$. Then $(b, [a, c]) = (b, b_k, [a, c]) = ((b, b_k), [a, c]) = [a, (b, b_k, c)] = [a, (b, c)]$ since $\mathfrak{A}_k$ is modular.

(2) $a, b \in \mathfrak{A}_k$, $c \in \mathfrak{S}_{k-1}$. Then $[a, (b, c)] = [a, (b, b_k, c)] = [a, (b, (b_k, c))] = (b, [a, (b_k, c)]) = (b, [[a, a_{k-1}], (b_k, c)]) = (b, b_k, [a, a_{k-1}, c]) = (b, [a, c])$ since $\mathfrak{A}_k$ and $\mathfrak{S}_{k-1}$ are modular.

(3) $a \in \mathfrak{A}_k$, $b, c \in \mathfrak{S}_{k-1}$. Then $(b, [a, c]) = (b, [a, [a_{k-1}, c]]) = [a, a_{k-1}, (b, c)] = [a, (b, c)]$ since $\mathfrak{S}_{k-1}$ is modular.

(4) $c \in \mathfrak{A}_k$, $a, b \in \mathfrak{A}_{k-1}$. Then $(b, [a, c]) = (b, [a, [a_{k-1}, c]]) = [a, (b, [a_{k-1}, c])] = [a, (b, b_k, [a_{k-1}, c])] = [a, a_{k-1}, (b, b_{k-1}, c)] = [a, (b, c)]$ since $\mathfrak{A}_k$ and $\mathfrak{S}_{k-1}$ are modular.

Hence in any case $[a, (b, c)] = (b, [a, c])$ and $\mathfrak{S}_k$ is thus modular.

THEOREM 4.2. *Let $\mathfrak{S}$ be a Birkhoff lattice. Then if $\mathfrak{S}_a$ contains a characteristic element which is not simple, there is an element $b \supset a$ such that $\mathfrak{S}_b$ is non-modular.*

Proof. Let c be a characteristic element of $\mathfrak{S}_a$ which is not simple. Then by Theorem 3.2 there is an element x and a complement c' such that $(c', x) > x$, $c' \not> a = [c', x]$. Let a_1 be the union of the points of $\mathfrak{S}_a$ which are divisible by $[u_a, x]$. Now if a_i has been defined, let a_{i+1} be the union of the points of $\mathfrak{S}_{a_i}$ which are divisible by $[u_{a_i}, x]$. Clearly $a_{i+1} \supset a_i$. Also a_{i+1} is distinct from a_i if $a_i \neq x$. Hence by the ascending chain condition $x = a_k$ for some k. Furthermore $a_{i+1} \in \mathfrak{S}_{a_i}$ since a_{i+1} is a union of points of $\mathfrak{S}_{a_i}$ and $u_{a_i} \in \mathfrak{S}_{a_{i+1}}$ since u_{a_i} is a union of points of $\mathfrak{S}_{a_{i+1}}$. Hence $\mathfrak{S}_{a_{i+1}}$ is an extension of $\mathfrak{S}_{a_i}$. Let $\mathfrak{S}_k$ be the set sum of $\mathfrak{S}_{a_1}, \cdots, \mathfrak{S}_{a_k}$. Then $\mathfrak{S}_k$ is a sublattice of $\mathfrak{S}$ and is non-modular since $(c', x) > x$ but $c' \not> [c', x]$ and $c', x \in \mathfrak{S}_k$. Hence by Lemma 4.2 $\mathfrak{S}_{a_j}$ is non-modular for some j. If we set $a_j = b$, the theorem follows.

Combining Theorems 4.1 and 4.2 we have

THEOREM 4.3. *Let $\mathfrak{S}$ be a Birkhoff lattice. Then the number of irreducible components is unique in the reduced decompositions of each element of $\mathfrak{S}$ if and only if $\mathfrak{S}_a$ is modular for each element a.*

Proof. The necessity follows immediately from Theorem 4.1. If each $\mathfrak{S}_a$ is modular, then by Theorem 4.2 each characteristic element is simple and hence the number of components is unique by Theorem 4.1.

If $\mathfrak{S}_a$ is modular, let x be any element of $\mathfrak{S}_a$. Furthermore, let $p_1, \cdots, p_k$ be a maximal union independent set of points of $\mathfrak{S}_a$ divisible by x. Imbed $p_1, \cdots, p_k$ in a maximal union independent set of points $p_1, \cdots, p_n$. Then $x = [x, (p_1, \cdots, p_n)] = (p_1, \cdots, p_k, [x, (p_{k+1}, \cdots, p_n)])$. If $[x, (p_{k+1}, \cdots, p_n)] \neq a$, then $[x, (p_{k+1}, \cdots, p_n)] \supset p > a$ and hence $(p_1, \cdots, p_k) \supset p$. But then $a = [(p_1, \cdots, p_k), (p_{k+1}, \cdots, p_n)] \supset p$, and this is impossible. Thus $x = (p_1, \cdots, p_k)$ and hence $\mathfrak{S}_a$ is a point lattice. The statement of Theorem 4.3 is thus equivalent to that given in the introduction.

Kurosch has shown that in a modular lattice if $a = [q_1, \cdots, q_k] = [q_1', \cdots, q_l']$ are two irreducible decompositions of a, then each q_i may be replaced by a q_i', and conversely, without changing the representation. The unicity of the number of components, of course, follows from this replacement property. Now

Theorem 4.1 shows that the converse is true for Birkhoff lattices, namely, that if the number of components in each reduced representation of a is unique, then the Kurosch replacement property holds for the reduced representations of a. For $\mathfrak{S}_a$ must be modular and have only simple characteristic elements by Theorem 4.1. But then the replacement property follows from the dual of Lemma 3.2 and Theorem 3.1.

COROLLARY 4.1. *Let $\mathfrak{S}$ be an exchange lattice (Birkhoff point lattice). Then the number of components in the reduced decompositions of the elements of $\mathfrak{S}$ is unique if and only if $\mathfrak{S}$ is modular.*

5. **Construction of Birkhoff lattices.** By definition u_a is never a characteristic element of $\mathfrak{S}_a$. Furthermore a is characteristic if and only if it is irreducible. The question then naturally arises whether there are any other elements of $\mathfrak{S}_a$ which can never be characteristic. In answer to this question we show that any $\mathfrak{S}_a$ may be imbedded in a Birkhoff lattice in which any given element of $\mathfrak{S}_a$ not equal to u_a or a is characteristic. The methods of construction are of considerable interest in themselves as they afford a means of obtaining Birkhoff lattices with various arithmetical properties.

Let us call two lattices $\mathfrak{S}_1$ and $\mathfrak{S}_2$ *compatible* if the set of common elements is a sublattice both of $\mathfrak{S}_1$ and $\mathfrak{S}_2$ and the operations of union and cross-cut in $\mathfrak{S}_1$ and $\mathfrak{S}_2$ agree in $\mathfrak{S}$.

LEMMA 5.1. *Let $\mathfrak{S}_1$ and $\mathfrak{S}_2$ be two compatible Birkhoff lattices satisfying the ascending chain condition. Moreover, let the common sublattice $\mathfrak{A}$ be complete in $\mathfrak{S}_1$. If a denotes the null element of $\mathfrak{S}_2$, let $a \epsilon \mathfrak{S}_1$ and let $x \supset a$, $x \epsilon \mathfrak{S}_1$ imply $x \epsilon \mathfrak{S}_2$. Then the set sum of $\mathfrak{S}_1$ and $\mathfrak{S}_2$ can be made into a Birkhoff lattice $\mathfrak{S}$ with $\mathfrak{S}_1$ and $\mathfrak{S}_2$ as sublattices.*

Proof. Let $\mathfrak{S}$ denote the set sum of $\mathfrak{S}_1$ and $\mathfrak{S}_2$. In $\mathfrak{S}$ we define division as follows: If $x \epsilon \mathfrak{S}_2$, then x divides y if and only if $x \supset (a, y)$, where the union is taken in $\mathfrak{S}_1$ or $\mathfrak{S}_2$ according as y is in $\mathfrak{S}_1$ or $\mathfrak{S}_2$. Since (a, y) is always in $\mathfrak{S}_2$, the division of the definition is in $\mathfrak{S}_2$. If $x \epsilon' \mathfrak{S}_2$, then x divides y if and only if $x \supset y$ in $\mathfrak{S}_1$. With respect to this division relation the union of two elements x and y is given by $((x, a), (y, a))$ if either x or y is in $\mathfrak{S}_2$ and is given by (x, y) in $\mathfrak{S}_1$ if $x, y \epsilon' \mathfrak{S}_2$. If $x \epsilon \mathfrak{S}_2$, let u_x be the union of the elements of $\mathfrak{A}$ which are divisible by x. The union always exists by the ascending chain condition. Then the cross-cut relation in $\mathfrak{S}$ is given by $[u_x, y]$ if $x \epsilon' \mathfrak{S}_1$, $y \epsilon' \mathfrak{S}_2$ and otherwise by $[x, y]$ where the cross-cut relation is in $\mathfrak{S}_1$ or $\mathfrak{S}_2$ according as both x and y are in $\mathfrak{S}_1$ or $\mathfrak{S}_2$. It follows immediately that the union and cross-cut relations in $\mathfrak{S}$ reduce to those in $\mathfrak{S}_1$ and $\mathfrak{S}_2$ for elements in $\mathfrak{S}_1$ and $\mathfrak{S}_2$ respectively.

Now let $x > y$ in $\mathfrak{S}$. If $y \epsilon \mathfrak{S}_2$, then $x \epsilon \mathfrak{S}_2$ and hence $x > y$ in $\mathfrak{S}_2$. If $y \epsilon' \mathfrak{S}_2$, then $x \epsilon \mathfrak{S}_1$. For if $x \epsilon' \mathfrak{S}_1$, then $x \supset a$ and $x \supset (y, a) \supset y$ in $\mathfrak{S}$. But then $x = (y, a)$ and $x \epsilon \mathfrak{S}_1$, and this is contrary to assumption. Thus if $y \epsilon \mathfrak{S}_1$, $x > y$ in $\mathfrak{S}_1$. Conversely, let $x > y$ in $\mathfrak{S}_2$. Then if $x \supset z \supset y$ in $\mathfrak{S}$, we have $z \epsilon \mathfrak{S}_2$ and $x = z$ or $z = y$. Hence $x > y$ in $\mathfrak{S}$. Let $x > y$ in $\mathfrak{S}_1$.

Suppose $x \supset z \supset y$ in $\mathfrak{S}$. If $z \epsilon \mathfrak{S}_1$, then $x = z$ or $z = y$ since $x > y$ in $\mathfrak{S}_1$. If $z \epsilon' \mathfrak{S}_1$, then $z \supset a$ and hence $x \supset a$. But then $x \supset (a, y) \supset y$ in $\mathfrak{S}_1$. If $x = (a, y)$, then $x = z$. If $(a, y) = y$, then $y \supset a$ and $y \epsilon \mathfrak{S}_2$. But then $x \epsilon \mathfrak{S}_2$ and $x, y \epsilon \mathfrak{A}$. Since $x > y$ in $\mathfrak{S}_1$ we have $x > y$ in $\mathfrak{A}$. Since $\mathfrak{A}$ is complete in $\mathfrak{S}_2$, we have $x > y$ in $\mathfrak{S}_2$. Thus $x = z$ or $z = y$. We conclude then that $x > y$ in $\mathfrak{S}$ if and only if $x > y$ in $\mathfrak{S}_1$ or $x > y$ in $\mathfrak{S}_2$.

Now let $x, y > [x, y]$ in $\mathfrak{S}$. If $[x, y] \epsilon \mathfrak{S}_2$, then $x, y \epsilon \mathfrak{S}_2$ and hence $x, y > [x, y]$ in $\mathfrak{S}_2$. But then $(x, y) > x, y$ in $\mathfrak{S}_2$ and hence in $\mathfrak{S}$ since $\mathfrak{S}_2$ is a Birkhoff lattice. If $[x, y] \epsilon' \mathfrak{S}_2$, then $x, y \epsilon \mathfrak{S}_1$. Hence $x, y > [x, y]$ in $\mathfrak{S}_1$. But then since $\mathfrak{S}_1$ is a Birkhoff lattice, we have $(x, y) > x, y$ in $\mathfrak{S}_1$ and thus in $\mathfrak{S}$. Hence $\mathfrak{S}$ is a Birkhoff lattice and the lemma is proved.

Lemma 5.2. *Let $\mathfrak{S}$ be a Birkhoff lattice satisfying the ascending chain condition and having a null element z. Let $\mathfrak{S}_p$ be the lattice obtained from $\mathfrak{S}$ by taking the direct sum of $\mathfrak{S}$ with a point p and deleting the simple elements. Then $\mathfrak{S}_p$ is a Birkhoff lattice containing $\mathfrak{S}$ as a complete sublattice.*

Proof. $\mathfrak{S}_p$ consists of two types of couples:

(1) $\{s, z\}$, $s \epsilon \mathfrak{S}$ and $s \neq u$,

(2) $\{s, p\}$, $s \epsilon \mathfrak{S}$, $u \not> s$.

If $\{x, y\}$ and $\{x_1, y_1\}$ are couples in $\mathfrak{S}_p$, we define $[\{x, y\}, \{x_1, y_1\}] = \{[x, x_1], [y, y_1]\}$. If $\{x, y\}$ and $\{x_1, y_1\}$ are $\mathfrak{S}_p$, then $\{[x, x_1], [y, y_1]\}$ is clearly in $\mathfrak{S}_p$. $(\{x, y\}, \{x_1, y_1\}) = \{(x, x_1), (y, y_1)\}$ if $\{(x, x_1), (y, y_1)\}$ is in $\mathfrak{S}_p$. Otherwise $(\{x, y\}, \{x_1, y_1\}) = \{u, p\}$. Then if $\{x, y\} \neq \{u, p\}$, $\{x, y\} > \{x_1, y_1\}$ if and only if $x > x_1$, $y = y_1$ or $x = x_1$, $y > y_1$. Hence the Birkhoff condition in $\mathfrak{S}_p$ follows from the Birkhoff condition in $\mathfrak{S}$. The correspondence $\{s, z\} \leftrightarrow s$, $s \neq u$ and $\{u, p\} \leftrightarrow u$ preserves union, cross-cut, and covering relations. Hence $\mathfrak{S}_p$ contains $\mathfrak{S}$ as a complete sublattice.

We shall refer to the process of Lemma 5.1 as the replacement of $\mathfrak{A}$ in $\mathfrak{S}_1$ by $\mathfrak{S}_2$. The process of Lemma 5.2 by which $\mathfrak{S}$ is imbedded in a lattice having the same unit and null elements will be called the imbedding of $\mathfrak{S}$ and $\mathfrak{S}_p$.

Lemma 5.3. *Let $\mathfrak{S}$ be an Archimedean Birkhoff lattice in which the unit element is the union of the points of $\mathfrak{S}$. Then $\mathfrak{S}$ may be imbedded in a Birkhoff lattice $\mathfrak{S}'$ having the same unit and null elements but having a simple element s and a chain of elements $s > s_1 > s_2 > \cdots > z$ all of which are union irreducible.*

Proof. Imbed $\mathfrak{S}$ in $\mathfrak{S}_{p_1}$. Then p_1 is union irreducible and $p_1 \epsilon' \mathfrak{S}$. Let $\mathfrak{S}_1$ be the quotient lattice of elements of $\mathfrak{S}_{p_1}$ which divide p_1. Replace $\mathfrak{S}_1$ by $\mathfrak{S}_{1p_2}$ in $\mathfrak{S}$ to give a Birkhoff lattice $\mathfrak{S}_2$. Now p_2 is union irreducible in $\mathfrak{S}_2$. For if $p_2 > x$, then $x \epsilon \mathfrak{S}_{1p_2}$ since $p_2 \epsilon' \mathfrak{S}_{p_1}$. But then $x = p_1$ since p_2 is a point of $\mathfrak{S}_{1p_2}$. Continuing in this manner, we get a chain of union irreducibles $z < p_1 < p_2 < \cdots < p_k$ which by the ascending chain condition must lead to a simple element $s = p_k$. $\mathfrak{S}' = \mathfrak{S}_k$ is the desired lattice.

We apply Lemma 5.3 to an arbitrary quotient lattice.

Theorem 5.1. *Let $\mathfrak{S}_a$ be the sublattice of a Birkhoff lattice $\mathfrak{S}$ belonging to the element a. Then if b is an arbitrary element of $\mathfrak{S}_a$ not the unit or null element,*

$\mathfrak{S}_a$ may be imbedded in a lattice $\mathfrak{S}'$ in which $\mathfrak{S}_a$ is the quotient lattice belonging to a and b is characteristic.

Proof. Let $\mathfrak{S}_b$ be the quotient lattice of elements of $\mathfrak{S}_a$ which divide b. Replace $\mathfrak{S}_b$ in $\mathfrak{S}_a$ by $\mathfrak{S}_b + b_1$, where b_1 is a point of $\mathfrak{S}_b + b_1$. Denote the lattice thus obtained by $\mathfrak{S}_{ab}$. Let $\mathfrak{S}_{b_1}$ be the quotient lattice of elements of $\mathfrak{S}_{ab}$ which divide b_1. Replace $\mathfrak{S}_{b_1}$ in $\mathfrak{S}_{ab}$ by $\mathfrak{S}'_{b_1}$ according to Lemma 5.3 to obtain a lattice $\mathfrak{S}'$. Then there is a chain of elements $b < b_1 < s_k < \cdots < s_1 < s$, where s is simple and $b_1, s_k, \cdots, s$ are union irreducible. Let b' be an arbitrary complement of b in $\mathfrak{S}_a$. Then $[b', s] = [b', s_1] = \cdots = [b', s_k] = [b', b_1] = [b, b'] = z$ and $(b', s) > s$. Hence b is characteristic by Theorem 3.1. Now let $\mathfrak{S}'_a$ be the quotient lattice associated with a in $\mathfrak{S}'$. Clearly $\mathfrak{S}_a \in \mathfrak{S}'_a$. If $x \in' \mathfrak{S}'_a$, then $x \supset b_1$ and $u_a \not\supset x$. Hence $x \in' \mathfrak{S}'_a$. Thus $\mathfrak{S}_a = \mathfrak{S}'_a$.

6. **Extension of reduced representations.** We turn now to the following problem: What are necessary and sufficient conditions that a given set of components of a may be extended to give a reduced decomposition of a? It is clearly necessary that the set of components be cross-cut independent. However, this is generally not sufficient if the elements of the lattice do not have unique irreducible decompositions. We obtain sufficient conditions by generalizing the notion of independence (Theorem 6.1). The above remark suggests the further problem of determining the conditions that a lattice must satisfy in order that every cross-cut independent set of components may be extended into a reduced decomposition. Theorem 6.2 gives a particularly simple answer, in case the number of components in the irreducible decompositions of the elements of the lattice is unique.

If $\mathfrak{S}_a$ is the quotient lattice associated with the element $a \in \mathfrak{S}$, let $\mathfrak{P}_a$ denote the elements of $\mathfrak{S}_a$ which can be expressed as a union of points of $\mathfrak{S}_a$. Now $\mathfrak{P}_a$ is clearly closed with respect to union and hence may be made into a lattice by defining cross-cut in terms of union. If c is a characteristic element of $\mathfrak{S}_a$, then there is exactly one characteristic element of $\mathfrak{S}_a$ which is associated with the same irreducibles as c and which belongs to $\mathfrak{P}_a$. Inasmuch as this characteristic element is the cross-cut of all characteristic elements associated with the same irreducibles as c, we shall call it the minimal characteristic element associated with c.

DEFINITION 6.1. A set of characteristic elements of $\mathfrak{S}_a$ is said to be sub-independent if the associated minimal characteristic elements are cross-cut independent in $\mathfrak{P}_a$.

We note that if the number of components is unique in $\mathfrak{S}$, then $\mathfrak{S}_a = \mathfrak{P}_a$ (Theorem 4.3) and sub-independence becomes ordinary cross-cut independence.

THEOREM 6.1. *Let $\mathfrak{S}$ be a Birkhoff lattice. Then a set $q_1, \cdots, q_k$ of components of a can be extended into a reduced decomposition of a if and only if the associated characteristic elements $c_1, \cdots, c_k$ of $\mathfrak{S}_a$ are sub-independent.*

Proof. Let $a = [q_1, \cdots, q_k, q_{k+1}, \cdots, q_n]$ be a reduced representation of a as a cross-cut of irreducibles. If $c_1, \cdots, c_k$ are not sub-independent, then with

suitable numbering $c_1 \supset \{c_2', \cdots, c_k'\}$, where c_i' is the minimal characteristic element associated with c_i and $\{\ ,\ \}$ denotes cross-cut in $\mathfrak{P}_a$. Let $[q_2, \cdots, q_n] \supset p > a$. Then $[c_2', \cdots, c_n'] \supset p$ and hence $\{c_2', \cdots, c_n'\} \supset p$. Then $c_1 \supset p$ and hence $a = [q_1, \cdots, q_n] \supset [c_1, \cdots, c_n] \supset \{c_1', \cdots, c_n'\} \supset p$. This contradicts $p > a$. Hence $c_1, \cdots, c_k$ are sub-independent.

Now let $c_1, \cdots, c_k$ be sub-independent. Let $p_1, \cdots, p_l$ be a maximal union independent set of elements covering a and divisible by $[q_1, \cdots, q_k]$ so that $\{c_1', \cdots, c_k'\} = (p_1, \cdots, p_l)$. Since $c_1, \cdots, c_k$ are sub-independent, there exist points $p_1', \cdots, p_k'$ such that $[c_1, \cdots, c_{i-1}, c_{i+1}, \cdots, c_k] \supset p_i'$, $c_i \not\supset p_i'$. Suppose $(p_2, \cdots, p_l, p_1', \cdots, p_k') \supset p_1$ say. Then since $(p_2, \cdots, p_l) \not\supset p_1$, there is a first i such that $(p_2, \cdots, p_l, p_1', \cdots, p_i') \supset p_1$. But then $(p_1, p_2, \cdots, p_l, p_1', \cdots, p_{i-1}') \supset p_i'$ by the Birkhoff condition. Hence with suitable numbering this case reduces to the case $(p_1, \cdots, p_l, p_2', \cdots, p_k') \supset p_1'$. But then $c_1 \supset (p_1, \cdots, p_l, p_2', \cdots, p_k') \supset p_1'$, and this contradicts $c_1 \not\supset p_1'$. Hence $p_1, \cdots, p_l, p_1', \cdots, p_k'$ are union independent and may be imbedded in a maximal union independent set $p_1, \cdots, p_l, p_1', \cdots, p_k', \cdots, p_n'$. Let $s_1 = (p_2, \cdots, p_n'), \cdots, s_l = (p_1, \cdots, p_{l-1}, p_1', \cdots, p_n')$. Now if $[c_1, \cdots, c_k, s_1, \cdots, s_l] \neq a$, we have $[c_1, \cdots, c_k, s_1, \cdots, s_l] \supset p > a$. Then since $[c_1, \cdots, c_k] \supset p$, it follows that $(p_1, \cdots, p_l) \supset p$. Hence $a = [(p_1, \cdots, p_l), s_1, \cdots, s_l] \supset p$. This is impossible since $p_1, \cdots, p_l, p_1', \cdots, p_n'$ generate a Boolean algebra. Thus $a = [c_1, \cdots, c_k, s_1, \cdots, s_l]$. We clearly cannot drop an s_i out of this representation since otherwise $a \supset p_i$ and this is impossible. Also we cannot drop out a c_i since then $c_i \supset p_i'$ and the definition of p_i' is contradicted. Hence $c_1, \cdots, c_k, s_1, \cdots, s_l$ is a cross-cut independent set of characteristic elements whose cross-cut is a. Thus by Theorem 3.1 there is a reduced representation of a containing $q_1, \cdots, q_k$.

COROLLARY 6.1. *Let $\mathfrak{S}$ be a Birkhoff lattice. Then an irreducible q is a component of a if and only if $q \supset a$ and $q \not\supset u_a$.*

COROLLARY 6.2. *Let $\mathfrak{S}$ be a Birkhoff lattice. Then if $q \supset b \supset a$ and q is a component of a, q is a component of b.*

For $u_b \supset u_a$. Hence if $q \supset u_b$, then $q \supset u_a$. This is impossible by Corollary 6.1.

COROLLARY 6.3. *If q and q' are components of a, then $q \supset q'$ implies $q = q'$.*

We are now ready to give a proof of Theorem 6.2 mentioned above using Theorem 6.1 and the Fundamental Lemma.

THEOREM 6.2. *Let $\mathfrak{S}$ be a Birkhoff lattice in which the number of components in the irreducible decompositions is unique for each element of $\mathfrak{S}$. Then each independent set of components of a can be extended into a reduced representation if and only if each characteristic element of $\mathfrak{S}_a$ belongs to exactly one irreducible component of a.*

Proof. Let $q_1, \cdots, q_k$ be an independent set of components of a. Let $c_1, \cdots, c_k$ be a set of characteristic elements associated with $q_1, \cdots, q_k$. We shall show that $c_1, \cdots, c_k$ are independent. If $c_1, \cdots, c_k$ are not independent,

then $c_1 \supset [c_2, \cdots, c_k]$, say. Now $[q_2, \cdots, q_k] \neq [c_2, \cdots, c_k]$, since otherwise $q_1 \supset c_1 \supset [q_2, \cdots, q_k]$ and the independence of $q_1, \cdots, q_k$ is contradicted. Let $[q_2, \cdots, q_k] \supset x_1 > [c_2, \cdots, c_k]$. Let $\mathfrak{A}_1$ be the sublattice of elements of $\mathfrak{S}_a$ contained between u_a and $[c_2, \cdots, c_k]$. Now $u_a \not\supset x_1$, since otherwise $c_i = [u_a, q_i] \supset [q_i, x_1]$ (Theorem 4.1) and hence $[c_2, \cdots, c_k] \supset [q_2, \cdots, q_k, x_1] = [q_2, \cdots, q_k]$. This is impossible. Hence by the Fundamental Lemma, $\mathfrak{A}_1$ and x_1 generate a sublattice which is the direct sum of $\mathfrak{A}_1$ and x_1. Let $\mathfrak{A}_2$ be the sublattice of elements of the form (y, x_1), where $y \in \mathfrak{A}$. Then $\mathfrak{A}_2$ is isomorphic to $\mathfrak{A}_1$ and is thus modular. The unit element of $\mathfrak{A}_2$ is (u_a, x_1) and the null element is x_1. Furthermore $c_1' = (c_1, x_1), \cdots, c_k' = (c_k, x_1)$ belong to $\mathfrak{A}_2$. Clearly $c_1' \supset (x_1, [c_2, \cdots, c_k]) = [(x_1, c_2), \cdots, (x_1, c_k)] = [c_2', \cdots, c_k']$. Now if $q_1 \not\supset (c_1, x_1)$, there is a second component q_1' of a such that $q_1' \supset (c_1, x_1)$, $q_1' \not\supset u_a$. But this contradicts the hypothesis of the theorem. Hence $q_1 \supset c_1'$. Now $[q_2, \cdots, q_k] \neq [c_2', \cdots, c_k']$ since otherwise $q_1 \supset c_1' \supset [q_2, \cdots, q_k]$. This is contrary to the independence of $q_1, \cdots, q_k$. Thus $[q_2, \cdots, q_k] \supset x_2 > [c_2', \cdots, c_k']$ for some x_2. We note that $c_i' = [q_i, (u_a, x_1)]$. For $(u_a, x_1) \supset [q_i, (u_a, x_1)] \supset c_i'$, and since $u_a > c_i$ we have $(u_a, x_1) > (c_i, x_1) = c_i'$. Hence either $c_i' = [q_i, (u_a, x_1)]$ or $q_i \supset (u_a, x_1) \supset u_a$. But $q_i \not\supset u_a$ and thus $c_i' = [q_i, (u_a, x_1)]$. Now $(u_a, x_1) \not\supset x_2$ since otherwise $c_i' = [q_i, (u_a, x_1)] \supset x_2$ and $[c_2', \cdots, c_k'] \supset x_2$, and this contradicts $x_2 > [c_2', \cdots, c_k']$. Hence by the Fundamental Lemma, the sublattice generated by $\mathfrak{A}_2$ and x_2 is the direct sum of $\mathfrak{A}_2$ and x_2. Continuing this process we get an infinite ascending chain $x_1 < x_2 < x_3 < \cdots$ and the ascending chain condition is contradicted. Hence if $q_1, \cdots, q_k$ are cross-cut independent, then $c_1, \cdots, c_k$ are also cross-cut independent. But now since $\mathfrak{P}_a = \mathfrak{S}_a$, $c_1, \cdots, c_k$ are also sub-independent. Thus by Theorem 6.1 $q_1, \cdots, q_k$ may be extended into a reduced representation.

If there is a characteristic element of $\mathfrak{S}_a$ which has two irreducibles to which it belongs, then the two irreducibles are cross-cut independent by Corollary 6.3. However, they cannot be extended into a reduced representation by Theorem 3.1. Hence the theorem follows.

Let $\mathfrak{S}$ be a Birkhoff lattice. Let c be a characteristic element of $\mathfrak{S}_a$. Then there is at least one characteristic element c_1 of $\mathfrak{S}_c$ which does not divide u_a. Similarly, there is at least one characteristic element c_2 of $\mathfrak{S}_{c_1}$ which does not divide u_a. Continuing in this manner we eventually get a chain $c \subset c_1 \subset c_2 \subset \cdots \subset c_k$, where c_k is an irreducible component of a and c_i is a characteristic element of $\mathfrak{S}_{c_{i-1}}$ such that $c_i \not\supset u_a$.

Theorem 6.2 may then be stated as follows:

THEOREM 6.3. *Let $\mathfrak{S}$ be a Birkhoff lattice in which the number of components is unique. Then each cross-cut independent set of components of a can be extended into an irreducible decomposition if and only if for each characteristic element c of $\mathfrak{S}_a$ there is exactly one chain $c \subset c_1 \subset c_2 \subset \cdots \subset c_k$, where c_k is an irreducible such that $c_k \not\supset u_a$ and c_i is a characteristic element of $\mathfrak{S}_{c_{i-1}}$. Moreover, in this case $\mathfrak{S}_{c_i} = \mathfrak{S}_{c_i}' + (c_i, u_a)$, where $\mathfrak{S}_{c_i}'$ consists of those elements of $\mathfrak{S}_{c_i}$ which do not divide u_a.*

Proof. If in any $\mathfrak{S}_{c_i}$ there are two characteristic elements c_{i+1} and c'_{i+1} which do not divide u_a, let $q_{i+1} \supset c_{i+1}$, $q_{i+1} \not\supset u_a$, $q'_{i+1} \supset c'_{i+1}$, $q'_{i+1} \not\supset u_a$. Then q_{i+1} and q'_{i+1} are distinct components of a which divide c and hence cannot be extended into an irreducible representation. The condition is sufficient since c_k is the single irreducible component of a which divides c.

Since $\mathfrak{S}_{c_i}$ is modular, to show that $\mathfrak{S}_{c_i} = \mathfrak{S}_{c_i} + (c_i, u_a)$ we have only to show that $x, y \not\supset u_a$, $x, y \in \mathfrak{S}_{c_i}$ implies $(x, y) \not\supset u_a$. But now c_{i+1} is the single simple element of $\mathfrak{S}_{c_i}$ which does not divide u_a. Let $\mathfrak{A}_i$ be the set of elements of $\mathfrak{S}_{c_i}$ which divide u_a. Then since $\mathfrak{S}_{c_i}$ is an atomic lattice $\mathfrak{S}_{c_i} = \mathfrak{A}_i \times c_{i+1}$ by the dual of the Fundamental Lemma. Then since $x, y \in \mathfrak{A}_i$, $c_{i+1} \supset x, y$ and hence $c_{i+1} \supset (x, y)$. Thus $(x, y) \not\supset u_a$. Hence $\mathfrak{S}_{c_i}$ is a sublattice of $\mathfrak{S}$ and the direct sum decomposition follows from the Fundamental Lemma.

COROLLARY 6.4. *Let $\mathfrak{S}$ be a modular lattice in which every complemented quotient lattice is irreducible. Then each cross-cut independent set of components of a can be extended into a reduced representation of a for all a if and only if $\mathfrak{S}$ is a chain of projective geometries connected by simple chains.*

The statement of Theorem 6.2 suggests the possibility of weakening the hypothesis that the number of components be unique. However, such a weakening will probably be very artificial since by using the methods of §5 the writer has constructed an example of a Birkhoff lattice $\mathfrak{S}$ in which each $\mathfrak{S}_a$ is a point lattice and each characteristic element is simple and has but one irreducible component as divisor. Furthermore, $\mathfrak{S}$ has only one element for which the number of components is not unique. Nevertheless not every cross-cut independent set of components can be extended into a reduced decomposition.

In §4 we pointed out that unicity of the number of components in a Birkhoff lattice implies the Kurosch replacement property. We conclude with a theorem which is a sort of converse result.

THEOREM 6.4. *Let $\mathfrak{S}$ be a Birkhoff lattice. Then if the Kurosch replacement property holds for the maximal representations of each element a, each representation of a is maximal, i.e., the number of components in the representations of a is unique.*

Proof. Let a be a simple element of $\mathfrak{S}_a$ which can be expressed as a union of points. Let b be any other element of $\mathfrak{S}_a$ which can be expressed as a union of points and such that $a \not\supset b$. Let $[a, b] \supset (p_1, \cdots, p_l)$, where $p_1, \cdots, p_l$ are a maximal union independent set of points of $\mathfrak{S}_a$ divisible by $[a, b]$. Let $a = (p_1, \cdots, p_l, p_{l+1}, \cdots, p_m)$ and $b = (p_1, \cdots, p_l, p'_{l+1}, \cdots, p'_k)$. Then there is a point of b, say p'_{l+1}, such that $p_1, \cdots, p_l, p_{l+1}, \cdots, p_m, p'_{l+1}$ form a maximal union independent set of points of $\mathfrak{S}_a$. Also $p_1, \cdots, p_l, p'_{l+1}, \cdots, p'_k$ can be imbedded in a maximal independent set $p_1, \cdots, p_l, p'_{l+1}, \cdots, p'_k, p'_{k+1}, \cdots, p'_r$. Let $a_1 = (p_2, \cdots, p_l, p_{l+1}, \cdots, p_m, p'_{l+1}), \cdots, a_{m+1} = (p_1, \cdots, p_l, p_{l+1}, \cdots, p_m)$. Similarly, let $a'_1 = (p_2, \cdots, p_l, p'_{l+1}, \cdots, p'_r), \cdots, a'_{m+1} = (p_1, \cdots, p_l, p'_{l+1}, \cdots, p'_{r-1})$. Now $[a'_1, \cdots, a'_l, a'_{l+2}, \cdots, a'_{m+1}] = p'_{l+1}$ by Lemma 3.2. Hence the only element of $a_1, \cdots, a_{m+1}$ which

can replace a'_{l+1} is a_{m+1}. But then $a'_1, \cdots, a'_l, a_{m+1}, a'_{l+2}, \cdots, a'_{m+1}$ must be simple elements of a Boolean algebra by Lemma 4.1. But $a = a_{m+1}$ and $b = [a'_{k+1}, \cdots, a'_{m+1}]$. Hence since a and b are elements of a Boolean algebra, we have $[a, b] = (p_1, \cdots, p_l)$ and $b > [a, b]$. Now let x and y be any two elements of $\mathfrak{S}_a$ which can be represented as a union of points of $\mathfrak{S}_a$. Then $x = [a_1, \cdots, a_k]$, where $a_1, \cdots, a_k$ are simple elements of $\mathfrak{S}_a$ which are unions of points. Hence $[x, y] = [a_1, \cdots, a_k, y]$ and hence $[x, y]$ is a union of points of $\mathfrak{S}_a$ by successive application of the result we have just obtained. But then the lattice generated by the points of $\mathfrak{S}_a$ is a modular sublattice of $\mathfrak{S}_a$ and hence is equal to $\mathfrak{S}_a$. $\mathfrak{S}_a$ is thus modular for each element a and the theorem follows from Theorem 4.3.

References

1. Garrett Birkhoff, *On the combination of subalgebras*, Proceedings of the Cambridge Philosophical Society, vol. 29(1933), pp. 441–464.
2. Garrett Birkhoff, *Abstract linear dependence and lattices*, American Journal of Mathematics, vol. 57(1935), pp. 800–804.
3. R. P. Dilworth, *Lattices with unique irreducible decompositions*, Annals of Mathematics, vol. 41(1940), pp. 771–777.
4. S. Mac Lane, *A lattice formulation for transcendence degrees and p-bases*, this Journal, vol. 4(1938), pp. 435–468.
5. O. Ore, *On the foundation of abstract algebra*, I, Annals of Mathematics, vol. 36(1935), pp. 406–437.
6. H. Whitney, *On the abstract properties of linear dependence*, American Journal of Mathematics, vol. 57(1935), pp. 509–533.

Yale University.

AN ANALOGUE OF GREEN'S THEOREM IN THE CALCULUS OF VARIATIONS

By Magnus R. Hestenes

1. **Introduction.** The purpose of the present paper is to study necessary and sufficient conditions for the equation

$$\iint_A (uz_x + vz_y + wz)\,dx\,dy = \int_C z(u\,dy - v\,dx) \tag{1}$$

to hold for every function $z(x, y)$ with continuous derivatives, where A is a region whose boundary C is composed of a finite number of rectifiable arcs without double points, every pair of which have at most an end point in common. This equation plays a fundamental rôle in the study of the properties of a minimizing surface for the double integral

$$\iint_A f(x, y, z, z_x, z_y)\,dx\,dy,$$

since the first variation of this integral is given by the first member of (1) with $u = f_{z_x}$, $v = f_{z_y}$, $w = f_z$. When w is continuous on $A + C$ and u, v have continuous derivatives on A with continuous limits on C, the criteria given below tell us that equation (1) holds if and only if $w = u_x + v_y$. Setting $z = 1$ in equation (1), one then obtains the Green's formula

$$\iint_A (u_x + v_y)\,dx\,dy = \int_C u\,dy - v\,dx.$$

On the other hand, equation (1) is an easy consequence of Green's formula when $w = u_x + v_y$. Our theorems therefore can be regarded as extensions of Green's theorem. Since our proofs are not based on Green's theorem, they can be regarded as an alternate proof of Green's theorem. The arguments here given are quite different from those given earlier by Haar,[1] Coral,[2] Bliss[3] and others.[4]

Received November 21, 1940; in revised form, January 29, 1941; presented to the American Mathematical Society, April 15, 1939 under the title *On the first necessary condition for minima of double integrals.*

[1] A. Haar, *Zur Variationsrechnung,* Abhandlungen aus dem mathematischen Seminar des Hamburgischen Universität, vol. 8(1930), p. 1.

[2] M. Coral, *On the necessary conditions for the minimum of a double integral,* this Journal, vol. 3(1937), pp. 585–592.

[3] G. A. Bliss, *The calculus of variations, multiple integrals,* Lectures delivered at the University of Chicago (mimeographed).

[4] See A. Huke, *An historical and critical study of the fundamental lemma of the calculus of variations, Contributions to the Calculus of Variations, 1930,* University of Chicago Press, 1931, pp. 45–160.

In §5 below it is shown, by an elementary argument, that a simply closed rectifiable curve C can be approached uniformly by a sequence of simply closed regular curves interior to C and having uniformly bounded lengths. This result is used in §7 to obtain an elementary proof of Cauchy's theorem for a function of a complex variable. In §6 there is given an interesting treatment of the one-dimensional analogue of equation (1).

2. **Hypotheses and first theorem.** Unless otherwise expressly stated it will be assumed that the set A under consideration is an open region whose boundary C is composed of a finite number of continuous arcs without double points, each of which has a continuously turning tangent and any pair of which have at most an end point in common. The end points of these arcs will be called *corner points* of C. A region A of this type will be called *admissible*. It will be understood that the line integrals along C will be taken in the positive sense of C with respect to A.

By an *admissible function* $z(x, y)$ will be meant one that is continuous and has continuous derivatives on a neighborhood of the set $A + C$. Other classes of admissible functions could also be used.

The functions $u(x, y)$ and $v(x, y)$ appearing in the integrals (1) are assumed to be continuous on $A + C$. The function $w(x, y)$ is supposed to be integrable.

The following interesting lemma will be useful:

LEMMA 1. *The formula* (1) *holds for every admissible function z if and only if for all except a finite number of points P on $A + C$ there is a neighborhood N of P such that the formula* (1) *holds for every admissible function z having $z = 0$ on $A - AN$.*

Consider first the case when to every point P of $A + C$ there corresponds a neighborhood N of P such that the formula (1) holds for every admissible function z having $z = 0$ on $A - AN$. Then $A + C$ can be covered by a finite set of circles $K_1, \cdots, K_m$ with centers $P_1, \cdots, P_m$ on $A + C$ and radii $r_1, \cdots, r_m$ such that for each integer i $(i \leqq m)$ the formula (1) holds for every admissible function z having $z = 0$ on $A - AK_i'$, where K_i' is the circle of radius $2r_i$ about P_i. Let $h(t)$ be a continuous function of a single variable t that has a continuous non-negative derivative and is such that $h(t) = 0$ when $t \leqq 1$ and $h(t) = 1$ when $t \geqq 2$. For example, the function $h(t)$ having $h(t) = 0$ $(t \leqq 1)$, $h(t) = 1$ $(t \geqq 2)$ and $h(t) = \exp[g(t)/(1 - t)]$, $g(t) = \exp[1/(t - 2)]$ has these properties. In fact it has continuous derivatives of all orders. Set $h_i(x, y) = h(d_i(x, y)/r_i)$, where $d_i(x, y)$ is the distance from (x, y) to P_i. We then have $h_i = 0$ on K_i and $h_i = 1$ on $A - AK_i'$. Given an admissible function z, the functions z_i defined by the equations $z_1 = (1 - h_1)z$, $z_i = h_1 \cdots h_{i-1}(1 - h_i)z$ $(1 < i \leqq m)$ are admissible and $z_i = 0$ on $A - AK_i'$. It follows from our hypotheses that $L(z_i) = 0$, where

$$L(z) = \iint_A (uz_x + vz_y + wz)\,dx\,dy - \int_C z(u\,dy - v\,dx). \tag{2}$$

Moreover, on $A + C$ we have $z_1 + \cdots + z_m = (1 - h_1 \cdots h_m)z = z$, the product $h_1 \cdots h_m$ being zero on $A + C$ since the spheres $K_1, \cdots, K_m$ cover $A + C$ and $h_i = 0$ on K_i. By virtue of the linearity of $L(z)$ we have therefore

$$L(z) = L(z_1 + \cdots + z_m) = L(z_1) + \cdots + L(z_m) = 0,$$

as was to be proved.

Consider next the case when to each point P of $A + C$, except possibly for a finite set of points $P_1, \cdots, P_m$, there is a neighborhood N of P such that $L(z) = 0$ for every admissible function z having $z = 0$ on $A - AN$. Let r be a positive constant less than one quarter of the distance between any pair of the points $P_1, \cdots, P_m$. Set $H(x, y, r) = h_1 \cdots h_m$, where h_i is defined relative to P_i as in the last paragraph with $r_i = r$. Then $H = 0$ on the r-neighborhood K of the points $P_1, \cdots, P_m$ and $H = 1$ on $A - AK'$, where K' is the $2r$-neighborhood of these points. Moreover,

$$(3) \qquad H_x = H_y = 0 \quad \text{on} \quad A - AK'; \qquad |H_x| < \frac{M}{r}, \qquad |H_y| < \frac{M}{r} \quad \text{on} \quad K',$$

where M is an upper bound of $h'(t)$. Given an admissible function z, the function $Z = Hz$ is admissible and $Z = 0$ on K. By the argument made in the last paragraph it is seen that $L(Z) = 0$. Hence

$$\iint_A (uz_x + vz_y + wz)H\,dx\,dy + \iint_A z(uH_x + vH_y)\,dx\,dy = \int_C Hz(u\,dy - v\,dx).$$

Since $H = 1$ on $A - AK'$, it is clear that, as r tends to zero, the first and last integrals in this equation have as their respective limits the first and second members of equation (1). The second integral tends to zero with r. This follows because of the relations (3). The absolute value of this integral does not exceed the value $(2MM'/r)m4\pi r^2 = 8MM'm\pi r$, where M' is an upper bound of the values $|zu|$, $|zv|$ on A. This completes the proof of Lemma 1.

The following theorem is a fundamental one in the theory of the first variation of double integrals of the calculus of variations.

THEOREM 1. *The formula* (1) *holds for every admissible function z if and only if it holds for every admissible function z having $z = 0$ on a neighborhood of the boundary C of A.*

It is to be understood that the neighborhood of C may vary with the choice of z. Suppose now that the criterion described in the theorem holds. Then to each point P of A there is a neighborhood N of P such that $L(z) = 0$ for every admissible function z having $z = 0$ on $A - AN$. In order to show that $L(z) = 0$ for every admissible function z it is sufficient, by Lemma 1, to show that for each point P of C, not a corner point of C, there is a neighborhood N of P such that $L(z) = 0$ for every admissible function z having $z = 0$ on $A - AN$. To this end we can assume that the boundary point P under consideration is the

origin and that the inner normal to C at P is the positive y-axis. Let N be a neighborhood of P defined by the inequalities $-e < x < e$, $-e < y < e$, where e is so small that the boundary of A in N is defined by an equation of the form $y = y(x)$ $(-e < x < e)$, the function $y(x)$ being a single-valued continuous function having a continuous derivative. Set

$$H(x, y, r) = h\left(\frac{y - y(x)}{r}\right),$$

where $h(t)$ is the function used in the proof of Lemma 1 above. Then $H = 0$ when $y \leqq y(x) + r$ and $H = 1$ when $y \geqq y(x) + 2r$. Consider now an admissible function z having $z = 0$ on $A - AN$. The function Z that is equal to Hz on N and identically zero elsewhere is admissible. It is zero at all points of N for which $y \leqq y(x) + r$ and hence identically zero on a neighborhood of the boundary C of A. It follows from our hypotheses that $L(Z) = 0$ and hence that

$$\iint_{AN} (uz_x + vz_y + wz)H\,dx\,dy = \iint_{AN} z(-uH_x - vH_y)\,dx\,dy.$$

Since $z = 0$ on $A - AN$ and $H = 1$ on N when $y \geqq y(x) + 2r$, the first member of this equation has the first member of equation (1) as its limit as r tends to zero. The second integral in this equation has the second member of equation (1) as its limit. To prove this we first observe that $H_x = H_y = 0$ when $y \geqq y(x) + 2r$. By taking r sufficiently small and setting $y = y(x) + rt$, it is found that this integral has the value

$$\int_{-e}^{e} (Uy' - V)\,dx,$$

where

$$U(x, r) = \int_0^2 z[x, y(x) + rt]u[x, y(x) + rt]h'(t)\,dt$$

and V is obtained from this formula by replacing u by v. By the use of the relations $h(2) = 1$, $h(0) = 0$, it is seen that $\lim_{r=0} U = zu$, and similarly that $\lim_{r=0} V = zv$. The integral under consideration accordingly has the limit

$$\int_{-e}^{e} z[x, y(x)]\{u[x, y(x)]y'(x) - v[x, y(x)]\}\,dx$$

which is identical with the second member of equation (1), since $z = 0$ on $A - AN$. This proves Theorem 1.

COROLLARY 1. *The equation* (1) *holds for every admissible function z if and only if to each point P of A there is a neighborhood N of P such that the equation* (1) *holds for every admissible function z having $z = 0$ on $A - AN$.*

For by an argument like that made in the proof of Lemma 1 it can be seen that the criterion given in this corollary is equivalent to the one given in Theorem 1.

From Theorem 1 and Corollary 1 we obtain

COROLLARY 2. *The equation* (1) *holds for every admissible function* z *if and only if for every admissible subregion* A' *of* A *the equation* (1), *with* A *replaced by* A', *holds for every admissible function* z.

In this corollary one can restrict the subregions A' so that their closures lie in A.

3. **Analogue of Green's theorem.** If in equation (1) we set $z = 1$, we find that the relation

$$\iint_A w\,dx\,dy = \int_C u\,dy - v\,dx \tag{4}$$

holds. This result together with those described in the corollaries of Theorem 1 suggests the following analogue of Green's theorem, which is useful in the study of the first variation of the double integrals of the calculus of variations.

THEOREM 2. *The equation* (1) *holds for every admissible function* z *if and only if the equation*

$$\iint_{A'} w\,dx\,dy = \int_{C'} u\,dy - v\,dx \tag{5}$$

is true for every square A' *whose closure is in* A *and whose sides are parallel to the axes, the curve* C' *being the boundary of* A'.

Let A' be a square in A at a distance $q > 0$ from the boundary C of A and whose sides are parallel to the axes. Since each point of A is contained in a square of this type, it is sufficient, by Corollary 1 to Theorem 1, to show that equation (1) holds for every admissible function z having $z = 0$ on $A - A'A$ when the criterion described in Theorem 2 is satisfied. To this end let

$$U(x, y, r) = \frac{1}{4r^2}\int_{x-r}^{x+r}\int_{y-r}^{y+r} u(s, t)\,ds\,dt \qquad (0 < r < q) \tag{6}$$

and let $V(x, y, r)$, $W(x, y, r)$ be the functions obtained from this formula when u is replaced by v, w respectively. By the use of the relations (5) and (6) it is found that the equation $W = U_x + V_y$ holds on $A' + C'$. If z is admissible, we have accordingly

$$\iint_{A'} (Uz_x + Vz_y + Wz)\,dx\,dy$$

$$= \iint_{A'} \{(Uz)_x + (Vz)_y\}\,dx\,dy = \int_{C'} z(U\,dy - V\,dx),$$

the last equality being obtained by an integration by parts. Taking the limit, as r tends to zero, we find that

$$\iint_{A'} (uz_x + vz_y + wz)\,dx\,dy = \int_{C'} z(u\,dy - v\,dx).$$

From this equation it is seen that equation (1) holds when $z = 0$ on $A - AA'$. This completes the proof of Theorem 2.

COROLLARY. *If equation* (5) *holds for every square A' whose closure is in A and whose sides are parallel to the axes, then equation* (4) *is also true.*

As a consequence of Theorem 2 we have the further

THEOREM 3. *Suppose that given any square $a \leqq x \leqq b$, $c \leqq y \leqq d$ in A the function $u(x, y)$ is absolutely continuous in x on $a \leqq x \leqq b$ for each y (almost everywhere) on $c \leqq y \leqq d$ and the function $v(x, y)$ is absolutely continuous n y on $c \leqq y \leqq d$ for each x (almost everywhere) on $a \leqq x \leqq b$. Suppose further that the derivatives u_x, v_y are integrable on A. Then the formula* (1) *holds for every admissible function z if and only if $w = u_x + v_y$ almost everywhere on A.*

For by iterated integration it is seen that the formula

$$\iint_{A'} (u_x + v_y)\,dx\,dy = \int_{C'} u\,dy - v\,dx$$

holds for every square A' in A whose sides are parallel to the axes and whose closure is in A. It follows from the last theorem that equation (1) holds when $w = u_x + v_y$ almost everywhere on A. On the other hand, if equation (1) is true, then the equation (5) is true for every square A' in A. We have accordingly the relation

$$\iint_{A'} (w - u_x - v_y)\,dx\,dy = 0$$

for every square A' in A. But this is true only if $w = u_x + v_y$ almost everywhere on A, as was to be proved.

COROLLARY (GREEN'S THEOREM). *If the functions u and v have the properties described in Theorem 3, the equation*

$$\iint_{A} (u_x + v_y)\,dx\,dy = \int_{C} u\,dy - v\,dx$$

holds.

4. **The case of rectifiable boundaries.** By a *regular curve* will be meant a continuous curve that can be divided into a finite number of subarcs on each of which it has a continuously turning tangent. The following lemma will be useful.

LEMMA 2. *Let C be a simply closed rectifiable curve of length l defined by the equations*

$$x = x(t), \qquad y = y(t) \qquad (0 \leqq t \leqq l), \tag{7}$$

where t denotes arc length. Let A be the set of points interior to C. There exists a sequence (C_n) of simply closed regular curves on $A + C$ defined by equations of the form

$$x = x_n(t), \qquad y = y_n(t) \qquad (0 \leqq t \leqq l;\, n = 1, 2, \cdots) \tag{8}$$

such that $\lim_{n=\infty} x_n(t) = x(t)$, $\lim_{n=\infty} y_n(t) = y(t)$ uniformly on $(0 \leqq t \leqq l)$ and such that absolute values of the derivatives $x_n'(t)$, $y_n'(t)$ are uniformly bounded with respect to n on $(0 \leqq t \leqq l)$. The lengths of the curves C_n are uniformly bounded.

Accepting for the moment the truth of this lemma, we can prove the following result:

THEOREM 4. *The results described in the last two sections are valid if we extend the definition of admissible sets A to include regions A whose boundary C is composed of a finite number of rectifiable arcs without double points, every pair of which have at most an end point in common.*

In order to prove this result we can suppose without loss of generality that the boundary C of A is a simply closed rectifiable curve C. Let $C_1, C_2, \cdots$ be regular curves related to C as described in Lemma 2 above and let A_n be the set of points interior to C_n. If the criterion described in Theorem 2 is satisfied, we have, by Theorem 2, the relations

$$\iint_{A_n} (uz_x + vz_y + wz)\, dx\, dy = \int_{C_n} z(u\, dy - v\, dx) \qquad (n = 1, 2, \cdots)$$

when z is admissible. As n becomes infinite, the first member of this equation has the first member of equation (1) as its limit. Similarly, the second member of this equation has as its limit the second member of equation (1), by virtue of a theorem given by Hobson.[5] An elementary proof of this fact can be made by an argument like that used at the end of §7. The conclusions described in Theorem 2 are therefore valid when C is made up of rectifiable arcs. Theorem 4 is now an easy consequence of this result and those described in the preceding sections.

5. **Proof of Lemma 2.** In order to prove Lemma 2 let A be the set of points interior to the given rectifiable curve C. Let O be a fixed point of A and choose a nearest point Q_0 of C to O. Set $r_n = e/2^{n+1}$, where e is the length of the line OQ_0. It will be shown below that, given an integer n, there exists a simply closed curve C_n on $A + C$ composed of circular arcs of radius r_n such that if $P_0, P_1, \cdots, P_m = P_0$ are the consecutive corners of C_n, the centers Q_1,

[5] E. W. Hobson, *Theory of Functions of a Real Variable*, vol. 2, p. 422.

$Q_2, \cdots, Q_m = Q_0$ of the circular arcs $P_0P_1, P_1P_2, \cdots, P_{m-1}P_m$ are consecutive points on C dividing C into m subarcs of lengths at least r_n. The point Q_0 is the one chosen above and is the same for all curves C_n.

If we grant for the moment the existence of curves C_n of the type just described, the proof of Lemma 2 can be made by setting up a correspondence between the curves C and C_n as follows: Let the point P_i on C_n correspond to the point Q_i on C. A point P on the circular arc $P_{i-1}P_i$ will be made to correspond to the point Q on the subarc $Q_{i-1}Q_i$ of C which divides this subarc in the same ratio relative to arc length as P divides $P_{i-1}P_i$, order being preserved. Assign to each point P on C_n the parameter value t belonging to the corresponding point Q in the equations (7) of C. One obtains thereby equations (8) for C_n. At a point P on $P_{i-1}P_i$ the absolute values of the derivatives $x_n'(t)$, $y_n'(t)$ cannot exceed the ratio of the arc lengths of $P_{i-1}P_i$ and $Q_{i-1}Q_i$ and hence are less than the value $2\pi r_n/r_n = 2\pi$. These derivatives are therefore uniformly bounded with respect to n. The length of C_n does not exceed the value $2\pi m r_n \leqq 2\pi l$. Moreover, given a constant $\epsilon > 0$, there is an integer n' such that when $n \geqq n'$ we have

$$(9) \qquad | x_n(t) - x(t) | < \epsilon, \qquad | y_n(t) - y(t) | < \epsilon.$$

To prove this we observe first that the distance between the points Q_{i-1} and Q_i corresponding to the points P_{i-1} and P_i of C_n does not exceed the length $2r_n$ of the polygon $Q_{i-1}P_{i-1}Q_i$. The distance between successive points in the sequence of points $Q_0, Q_1, \cdots, Q_m = Q_0$ on C determined by C_n is therefore at most $2r_n$. Since C is rectifiable and $\lim_{n=\infty} r_n = 0$, it follows that there is an integer n' such that when $n \geqq n'$ the points $Q_0, \cdots, Q_m$ described above divide C into subarcs of lengths at most $\frac{1}{2}\epsilon$. Increase n' so that $3r_n < \frac{1}{2}\epsilon$ when $n \geqq n'$. Let P be a point on the subarc $P_{i-1}P_i$ of C_n ($n \geqq n'$) and let Q be the corresponding point on the subarc $Q_{i-1}Q_i$ of C. The distance from P to Q does not exceed the length of the polygon $PP_{i-1}Q_{i-1}Q$ and hence does not exceed the value $2r_n + r_n + \frac{1}{2}\epsilon < \epsilon$. From this result it is seen that the inequalities (9) hold when $n \geqq n'$.

The proof of Lemma 2 will be complete if we show the existence of the curve C_n described at the beginning of this section. This can be done in a number of ways. We shall proceed as follows: Let P be the point on the line OQ_0, described above, at the distance r_n from Q_0. With Q_0 as the center draw through P the shortest circular arc K_0 on $A + C$ cutting C in points R and R_0. Denote by C' the subarc R_0R of C not containing Q_0 and let the points of C' be ordered so that R_0 is its initial end point and R is its final end point. Since the distance from P to C' exceeds r_n, there is at least one point Q on C', namely $Q = R_0$, with the property that there exists on $A + C$ a circular arc of radius r_n about Q joining a point on the subarc PR_0 of K_0 to a point on the subarc QR of C'. Moreover, it is easily seen that there is a last point Q_1 on C' having this property. Let K_1 be the shortest circular arc on $A + C$ of radius r_n and with center at Q_1 that joins a point P_0 on PR_0 to a point R_1 on the subarc Q_1R

of C'. The point P_0 is the only point of K_1 on PR_0 and R_1 is the only point of K_1 on the subarc Q_1R of C', since otherwise K_1 could be shortened. Furthermore, the length of K_1 exceeds $\frac{1}{3}\pi r_n$. If this were not the case, one could draw on $A + C$ a circular arc about R_1 of radius r_n joining a point of PR_0 to a point on the subarc R_1R of C', contrary to our choice of Q_1. Since the length of K_1 exceeds $\frac{1}{3}\pi r_n$, there is a point Q on the subarc R_1R of C', namely $Q = R_1$, for which there exists on $A + C$ a circular arc about Q of radius r_n joining a point of K_1 to a point on the subarc QR of C'. Let Q_2 be the last point Q on C' having this property and draw about Q_2 on $A + C$ the shortest circular arc K_2 of radius r_n joining a point P_1 of K_1 to a point R_2 on the subarc Q_2R of C'. The point P_1 is the only point of K_2 on K_1 and R_2 is the only point of K_2 on Q_2R. The length of K_2 exceeds $\frac{1}{3}\pi r_n$, since otherwise our choice of Q_2 or Q_1 could be contradicted. Moreover, the arc K_2 has no point in common with the subarc PR_0 of K_0 in view of our choice of Q_1. Proceeding in this manner, one obtains on $A + C$ circular arcs $K_1, K_2, \cdots, K_{m-2}$ of radius r_n with centers Q_1, $Q_2, \cdots, Q_{m-2}$ on C'. The arc K_i is the shortest circular arc on $A + C$ of radius r_n with center at Q_i which joins a point P_{i-1} on K_{i-1} to a point R_i on the subarc Q_iR of C', and Q_i is the last point on the subarc $R_{i-1}R$ of C' related to K_{i-1} in this manner. The point P_{i-1} is the only point of K_i on K_{i-1}, and R_i is the only point of K_i on Q_iR. The length of K_i exceeds $\frac{1}{3}\pi r_n$. In view of our choice of Q_i and Q_1, no arc K_j $(j > i)$ has a point in common with K_{i-1} or with the subarc PR_0 of K_0. Finally, the length of the subarc $Q_{i-1}Q_i$ of C' is at least r_n, since this subarc contains the point R_{i-1} whose distance from Q_{i-1} is equal to r_n. In view of this fact and the rectifiability of C' the construction just described cannot be carried out indefinitely. There is accordingly a first integer m such that, the circular arc K_{m-2} joining K_{m-3} to a point R_{m-2} on the subarc $Q_{m-2}R$ of C' having been constructed, there is a point Q_{m-1} on the subarc $R_{m-2}R$ of C' for which there exists on $A + C$ a circular arc K_{m-1} about Q_{m-1} of radius r_n joining a point P_{m-2} on K_{m-2} to a point P_{m-1} on the subarc PR of K_0. Select K_{m-1} to be the shortest circular arc related to Q_{m-1} in this manner. Set $K_m = K_0$, $P_m = P_0$, $Q_m = Q_0$. The subarcs P_0P_1 of K_1, P_1P_2 of $K_2, \cdots, P_{m-1}P_m$ of K_m define a simply closed curve C_n on $A + C$ having the properties described at the beginning of this section. This completes the proof of Lemma 2.

COROLLARY. *The curves C_n described in Lemma 2 can be chosen to be interior to C.*

For let C_n be constructed as described above and choose a number r_n' on $r_n < r < r_n + 1/n$. If r_n' is chosen sufficiently near r_n, there is a simply closed curve C_n' interior to C_n composed of circular arcs of radius r_n' whose centers are at the points $Q_0, Q_1, \cdots, Q_{m-1}$ on C associated with C_n. The curve C_n can be replaced by C_n' in the above discussion, provided we also replace r_n by r_n'. The curve C_n', being interior to C_n, is interior to C. This proves the corollary.

COROLLARY. *The curves C_n described in Lemma 2 can be chosen to be polygons interior to C with sides parallel to the coördinate axes.*

For let C_n and C'_n be the curves related to C as described in the proof of the last corollary. Denote by P'_i the corner point of C'_n corresponding to the corner P_i of C_n. It is easily seen that there is a polygon C''_n with sides parallel to the axes in the region bounded by C_n and C'_n cutting the line $P_iP'_i$ in a single point P''_i such that the length of the subarc $P''_{i-1}P''_i$ of C''_n does not exceed $8r'_n$, where r'_n is the number used in the definition of the curve C'_n. Parameterize C''_n as described in the second paragraph of this section, the points P''_i playing the rôle of the points P_i. Observing that the distance from Q_i to P''_i (and to P''_{i-1}) is less than r'_n, one can show, by an argument like that used in the second paragraph of this section, that the curves C''_n have the properties described in Lemma 2.

The result described in the last corollary can be used to give a second proof of Green's theorem, by first showing that Green's theorem holds for simply closed polygons whose sides are parallel to the axes. This method has been used recently by Reid,[6] who, however, uses a weaker convergence theorem than the one here described. Similar methods have been used by Van Vleck,[7] Bray[8] and others.[9]

6. **The case of simple integrals.** The method used in the preceding pages leads to an interesting proof of a fundamental lemma for simple integral problems in the calculus of variations. Let $u(x)$, $v(x)$ be piecewise continuous functions on x_1x_2; that is, the interval x_1x_2 can be subdivided into a finite number of subintervals on each of which u and v are continuous. By an *admissible function* $z(x)$ will be meant one that is continuous and has a piecewise continuous derivative on x_1x_2.

THEOREM 5. *The equation*

$$\int_{x_1}^{x_2} (uz + vz')\,dx = v(x_2)z(x_2) - v(x_1)z(x_1) \tag{10}$$

holds for every admissible function z if and only if it holds for every admissible function z vanishing at x_1 and x_2. In fact, equation (10) *holds for every admissible function z if and only if one has on x_1x_2 the relation*

$$v(x) = \int_{x_1}^{x} u(t)\,dt + v(x_1). \tag{11}$$

[6] W. T. Reid, *Green's lemma and related results*, to appear in the American Journal of Mathematics, vol. 63(1941).

[7] E. B. Van Vleck, *An extension of Green's lemma to the case of a rectifiable boundary*, Annals of Mathematics, vol. 22(1921), pp. 226–237.

[8] H. E. Bray, *Green's lemma*, Annals of Mathematics, vol. 26(1925), pp. 278–286.

[9] Cf. K. Menger, *On Green's formula*, Proceedings of the National Academy of Sciences, vol. 26(1940), pp. 660–664.

In order to prove[10] this result let $h(x)$ be a function such that

$$h(x) = \frac{x - x_1}{r} \quad (x_1 \leqq x \leqq x_1 + r), \qquad h(x) = \frac{x_2 - x}{r} \quad (x_2 - r \leqq x \leqq x_2)$$

and $h(x) = 1$ elsewhere. If z is an admissible function, then $y = hz$ is an admissible function vanishing at x_1 and x_2. Under the hypothesis that equation (10) holds for every admissible function z vanishing at x_1 and x_2, we find upon replacing an arbitrary admissible function z by hz in (10) that

$$\int_{x_1}^{x_2} (uz + vz')h\,dx = \frac{1}{r}\int_{x_2-r}^{x_2} vz\,dx - \frac{1}{r}\int_{x_1}^{x_1+r} vz\,dx.$$

Letting r approach zero, one obtains equation (10) as a limit. This proves the first statement in the theorem. Setting $z = 1$ in (10), one obtains equation (11) with $x = x_2$. Moreover, if equation (10) holds for every admissible function z, it holds when x_2 is replaced by any value x' on x_1x_2, by virtue of the first part of the theorem applied to the interval x_1x'. Hence equation (11) holds on x_1x_2. Conversely, equation (10) follows from equation (11), since $v' = u$ on each subinterval of x_1x_2 on which u is continuous. This completes the proof of Theorem 5.

The fundamental lemma in the calculus of variations states that if the equation

$$\int_{x_1}^{x_2} vz'\,dx = 0$$

holds for every admissible function z vanishing at x_1 and x_2, then v is a constant. This result follows from equation (11) with $u = 0$. On the other hand, Theorem 5 is an easy consequence of this result and hence is equivalent to it. Of the two, it appears that Theorem 5 is the easier to apply in the study of the first variation of simple integrals.

7. **Cauchy's theorem.** It was shown in §5, by an elementary argument, that given a simply closed rectifiable curve C in the complex plane defined by functions $z(t) = x(t) + iy(t)$ $(0 \leqq t \leqq l)$ with arc length as parameter, there exist simply closed regular curves $C_1, C_2, \cdots$ interior to C which are of uniformly bounded lengths and are defined by functions $z_n(t) = x_n(t) + iy_n(t)$ $(0 \leqq t \leqq l;\ n = 1, 2, \cdots)$ that converge uniformly to $z(t)$ on $0 \leqq t \leqq l$. This fact can be used to give an elementary proof of the following well-known result:

THEOREM 6 (CAUCHY'S THEOREM). *Let C be a simply closed rectifiable curve. If $f(z)$ is holomorphic in C and continuous in and on C, then* $\int_C f(z)\,dz = 0$.

For let $C_1, C_2, \cdots$ be regular curves related to C as described above. Since $f(z)$ is holomorphic in and on C_n, one has $\int_{C_n} f(z)\,dz = 0$, by the usual proofs.

[10] If in equations (12) below one sets $z = 1$ and replaces x_2 by x, one obtains a proof of equations (11) communicated to me by Professor L. M. Graves.

By the use of the well-known result

$$\lim_{n \to \infty} \int_{C_n} f(z)\, dz = \int_C f(z)\, dz \tag{12}$$

it follows that $\int_C f(z)\, dz = 0$, as was to be proved.

An elementary proof of (12) can be made as follows: Let L be a bound for the length of the curves $C, C_1, C_2, \cdots$ and choose a constant $\epsilon > 0$. Since $f(z)$ is uniformly continuous in and on C, there is a constant $\delta > 0$ such that

$$|f(z) - f(z^*)| < \frac{\epsilon}{4L} \tag{13}$$

for every pair of points z, z^* in and on C for which $|z - z^*| < 4\delta$. Divide the interval $0 \leqq t \leqq l$ into m subintervals of lengths at most δ by points $t_0 = 0 < t_1 < t_2 < \cdots < t_m = l$. Choose n' so large that $|z(t) - z_n(t)| < \delta$ on $0 \leqq t \leqq l$ when $n \geqq n'$. The points $z^i = z(t_i)$, $z_n^i = z_n(t_i)$ $(n \geqq n';\ i = 0, 1, \cdots, m)$ then satisfy the inequalities $|z^i - z^{i-1}| < \delta$, $|z_n^i - z_n^{i-1}| < 3\delta$. One has accordingly, by (13), the inequalities

$$\left| \int_C f(z)\, dz - S \right| \leqq \tfrac{1}{4}\epsilon, \qquad \left| \int_{C_n} f(z)\, dz - S_n \right| \leqq \tfrac{1}{4}\epsilon \qquad (n \geqq n'), \tag{14}$$

where

$$S = \sum_{i=1}^{m} f(z^i)(z^i - z^{i-1}), \qquad S_n = \sum_{i=1}^{m} f(z_n^i)(z_n^i - z_n^{i-1}).$$

But

$$S - S_n = \sum_{i=1}^{m} [f(z^i) - f(z_n^i)](z_n^i - z_n^{i-1}) + \sum_{i=1}^{m} f(z^i)[(z^i - z_n^i) - (z^{i-1} - z_n^{i-1})].$$

Suppose $M = \sum_{i=1}^{m} |f(z^i)|$ and increase n' so that $|z(t) - z_n(t)| < \epsilon/8M$ on $0 \leqq t \leqq l$ when $n \geqq n'$. One has accordingly, by (13) and the formula for $S - S_n$, the inequality

$$|S - S_n| < \frac{\epsilon}{4L}\cdot L + M2\cdot\frac{\epsilon}{8M} = \tfrac{1}{2}\epsilon \qquad (n \geqq n'). \tag{15}$$

Combining the relations (14) and (15), one obtains the inequality

$$\left| \int_C f(z)\, dz - \int_{C_n} f(z)\, dz \right| < \epsilon \qquad (n \geqq n').$$

The limit (12) is thereby established.

THE UNIVERSITY OF CHICAGO.

A GENERALIZATION OF THE AUMANN-CARATHÉODORY "STARRHEITSSATZ"

By Maurice H. Heins

1. **Introduction.** The "Starrheitssatz" of Aumann and Carathéodory [2][1] may be stated as follows:

Let G_w be a multiply-connected region of the w-plane, the boundary of which contains at least three points, and let $W = W(w)$ be analytic and single-valued for $w \in G_w$. Further let $W = W(w)$ satisfy the requirements:

(i) *there exists a $\zeta \in G_w$ such that $W(\zeta) = \zeta$,*

(ii) *$w \in G_w$ implies $W(w) \in G_w$.*

Then there exists a positive constant $\Omega(\zeta, G_w)$ less than unity such that, if $W = W(w)$ is not a (1, 1) map of G_z onto itself, then $| W'(\zeta) | \leqq \Omega(\zeta, G_w)$.

If we denote by C_1 the class of (1, 1) conformal maps of G_w onto itself, and by C_2 the class of all other single-valued maps which are analytic for $w \in G_w$ and have their images in G_w, then the "Starrheitssatz" asserts that there exists no sequence of maps $\{W_n(w)\}$ of C_2 with $W_n(\zeta) = \zeta$ $(n = 1, 2, \cdots)$ which converges continuously to a map of class C_1 for $w \in G_w$. Conversely, if we can establish that no map of C_1 can be expressed as the limit of a sequence of maps of class C_2, then the "Starrheitssatz" follows immediately. This results from the fact that, if $W = W_0(w)$ is a map of either class C_1 or C_2 with the properties

(i) $W_0(\zeta) = \zeta$,

(ii) $| W_0'(\zeta) | = 1$,

then $W_0(w)$ is necessarily a member of class C_1.

The "Starrheitssatz" is restrictive in its hypotheses. It requires that $\zeta \in G_w$ be a fixed point of the maps considered. It is therefore natural to seek a generalization of the "Starrheitssatz" which does not make such stringent requirements on the class of maps considered. The alleged proposition that no map of C_1 can be expressed as the limit of a sequence of maps of C_2 offers such a generalization. In this paper we shall establish a proposition of this type.

We need not restrict our attention to plane regions G_w. Instead we may very well consider abstract Riemann surfaces[2] F_w and single-valued conformal maps $W = W(w)$ of F_w into itself. We shall require that F_w be not simply-connected, that $\bar{F}_w$, the *universal covering surface* of F_w, be of hyperbolic type. Let $w = w(z)$ denote any conformal uniformizing mapping which defines $| z | < 1$ as a smooth, unbounded covering surface of F_w. The map $w = w(z)$ is automorphic under a group $\mathfrak{G}$ of linear fractional transformations T which map $| z | < 1$ onto

Received December 12, 1940.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

[2] See [6]. We adopt the notation and definitions of this text.

itself. We shall further require that the points on $|z| = 1$ where $\mathfrak{G}$ ceases to be properly discontinuous constitute an infinite set.

Under these circumstances, we shall establish that *there exists no sequence of maps* $W_n(w)$ $(\not\equiv w)$ $(n = 1, 2, \cdots)$ *which converges pointwise to* w *for* $w \in F_w$.

This theorem implies several immediate corollaries: (i) the theorem of Klein and Poincaré ([6], pp. 163–165): namely, that under the hypotheses of our theorem the group of (1, 1) conformal maps of F_w onto itself is properly discontinuous; hence, (ii) the theorems relating to the number of (1, 1) conformal maps of a closed Riemann surface or an open planar Riemann surface of finite connectivity onto itself; (iii) the "Starrheitssatz" already mentioned; (iv) the theorem that no (1, 1) conformal map of F_w onto itself, $W_0(w)$, can be expressed as the limit of a sequence $\{W_n(w)\}$, where $W_n(w) \not\equiv W_0(w)$ $(n = 1, 2, \cdots)$. Corollary (iv) is an immediate consequence of our principal theorem and requires no further comment.

2. **A preliminary lemma.**

In this section we shall establish a simple lemma which is the basis of the proof of our theorem.

LEMMA 2.1. *Let S denote a hyperbolic transformation of $|z| < 1$ onto itself. Then the only functions $\varphi(z)$ which are analytic and of modulus less than unity for $|z| < 1$ and which satisfy the functional relation*

$$\varphi(S) = S[\varphi(z)] \tag{2.1}$$

are the hyperbolic transformations of $|z| < 1$ onto itself with the same fixed points as S.

We may map $|z| < 1$ one to one and conformally onto $\Re[x] > 0$ in such a manner that the fixed points of S correspond to 0 and ∞ in the x-plane. Call this map V. Consider the transform of $\varphi(z)$ with respect to V, $V[\varphi(V^{-1}x)]$, denoting it by $\psi(x)$. It is clear that $\psi(x)$ is analytic and has the property that $\Re[\psi(x)] > 0$ for $\Re[x] > 0$. Further, $\psi(x)$ satisfies a functional relation of the form

$$\psi(\lambda x) = \lambda\psi(x), \tag{2.2}$$

where λ is a positive constant greater than unity.

Let us consider the implications of (2.2) for $\psi(x)$. Let c $(\geqq 0)$ denote the angular derivative at infinity ([5], pp. 52–55) of $\psi(x)$. c can be calculated from the relation

$$\lim_{n\to\infty} \frac{\psi(\lambda^n x_0)}{\lambda^n x_0} = c, \tag{2.3}$$

where x_0 is an arbitrary point of $\Re[x] > 0$. But (2.2) implies that

$$\psi(\lambda^n x) = \lambda^n \psi(x);$$

hence

$$\lim_{n\to\infty} \frac{\lambda^n \psi(x_0)}{\lambda^n x_0} = c,$$

or

$$\frac{\psi(x_0)}{x_0} = c.$$

We infer from this last relation that c is positive. Now x_0 is an arbitrary point of $\Re[x] > 0$; hence $\psi(x)$ is of the form cx. Conversely, every function of the form cx $(c > 0)$ satisfies (2.2) and has a positive real part for $\Re[x] > 0$. Returning to $|z| < 1$ we infer immediately the truth of Lemma 2.1.

3. **The principal theorem.** Lemma 2.1, together with the facts that $\mathfrak{G}$ is properly discontinuous in $|z| < 1$ and that, under the hypotheses of our theorem, $\mathfrak{G}$ always contains at least one hyperbolic transformation ([1], p. 62), furnishes an immediate proof of the theorem which we wish to establish.

Let $\varphi_n(z)$ denote a *transform* of $W_n(w)$ with respect to the uniformization mapping $w(z)$. That is, we define $\varphi_n(z)$ up to a linear fractional transformation of $\mathfrak{G}$ by the relation

$$w(\varphi_n(z)) = W_n(w(z)) \qquad (n = 1, 2, \cdots). \tag{3.1}$$

So defined, $\varphi_n(z)$ is analytic and of modulus less than unity for $|z| < 1$. Further, $\varphi_n(z)$ satisfies a system of functional equations

$$\varphi_n(T) = U_T^{(n)}[\varphi_n(z)] \qquad (T, U_T^{(n)} \in \mathfrak{G}; n = 1, 2, \cdots). \tag{3.2}$$

These relations follow immediately from the fact that, for $T \in \mathfrak{G}$, $\varphi_n(T)$ is a transform of $W_n(w)$ whenever $\varphi_n(z)$ is.

Assuming, contrary to the assertion of the theorem which we wish to establish, that there exists a sequence $\{W_n(w)\}$ which converges pointwise to w for $w \in F_w$, we may choose z itself as the transform of w and the transforms of $W_n(w)$, $\varphi_n(z)$ $(n = 1, 2, \cdots)$ in such a manner that the relation

$$\lim_{n\to\infty} \varphi_n(z) = z \tag{3.3}$$

obtains pointwise and hence continuously for $|z| < 1$.

We have remarked that the group $\mathfrak{G}$ contains a hyperbolic transformation S_1. Let S_2 denote any transformation of $\mathfrak{G}$ with some fixed point distinct from the fixed points of S_1. The group $\mathfrak{G}$ is properly discontinuous. Hence for $T = S_k$ $(k = 1, 2)$ the relations (3.2) and (3.3) imply that for $n > n_0$

$$U_{S_1}^{(n)} = S_1; \qquad U_{S_2}^{(n)} = S_2. \tag{3.4}$$

By virtue of Lemma 2.1, $\varphi_n(z)$ is a hyperbolic transformation of $|z| < 1$ onto itself with the same fixed points as S_1 for $n > n_0$. But for these same values of n, $\varphi_n(z)$ also satisfies the relation

$$\varphi_n(S_2) = S_2[\varphi_n(z)] \qquad (n > n_0). \tag{3.5}$$

The function $\varphi_n(z)$ is a hyperbolic transformation with the same fixed points as S_1. The relation (3.5) implies that, if $\varphi_n(z) \not\equiv z$, S_2 is hyperbolic and has the same fixed points as $\varphi_n(z)$. This is contrary to the choice of S_2. Equation (3.1) implies for $n > n_0$ that

$$W_n(w) \equiv w.$$

Hence we have

THEOREM 3.1. *If F_w is a Riemann surface, which is not simply-connected and which has the properties that $\tilde{F}_w$ is of hyperbolic type and that the fundamental group associated with F_w is not cyclic, then the identical map of F_w onto itself can never be expressed as the limit of a sequence $\{W_n(w)\}$ of single-valued conformal maps of F_w onto itself, where $W_n(w) \not\equiv w$ $(n = 1, 2, \cdots)$.*

In particular, we may infer that the group of (1, 1) conformal maps of F_w onto itself is properly discontinuous ([6], pp. 163–165).

Furthermore, the theorem that the number of (1, 1) conformal maps of a plane region G_w bounded by p (> 2) Jordan curves onto itself is finite, also follows [4]. For, if there were an infinite number of such maps, there would exist a sequence $\{W_n(w)\}$ of such maps converging continuously for $w \in G_w$ to a point ω of the boundary of G_w. By the continuity properties of $w = w(z)$, which maps $|z| < 1$ one to one and conformally onto $\tilde{G}_w$, we may associate with ω a point ζ of $|z| = 1$ with the property that

$$\lim_{z \to \zeta} w(z) = \omega \qquad (|z| < 1).$$

Hence we may assign to each $W_n(w)$ a transform $\varphi_n(z)$ $(n = 1, 2, \cdots)$ such that the sequence $\{\varphi_n(z)\}$ converges continuously to ζ as $n \to \infty$ for $|z| < 1$. It is well known that the group $\mathfrak{G}$ associated with $w(z)$ is properly discontinuous at ζ. Let T be any member of $\mathfrak{G}$ distinct from the identity. Then for $n > N(T)$, $\varphi_n(T) = \varphi_n(z)$. But all $\varphi_n(z)$ are linear fractional transformations of $|z| < 1$ onto itself, hence not automorphic. Therefore the number of (1, 1) conformal maps of G_w onto itself is finite.

The "Starrheitssatz" for regions of connectivity greater than two also follows from our preliminary remarks.

Theorem 3.1 is not true for Riemann surfaces which are not of the type described in the hypotheses of that theorem. One need only consider the rotations $W_\theta = e^{i\theta}w$ (θ real) of the annulus $\mathfrak{A}$: $r_1 < |w| < r_2$ $(r_1 > 0)$ to see that Theorem 3.1 is no longer valid. However, it is possible to replace Theorem 3.1 in this case by the proposition that for $\mathfrak{A}$ no map of class C_1 can be expressed as the limit of a sequence of maps of class C_2 (cf. §1).

The proof is easy. By studying the transforms of the maps $W(w)$ with respect to $w(z)$ which map $\tilde{\mathfrak{A}}$ one to one and conformally onto $|z| < 1$, we find that the transforms of the maps of class C_2 are automorphic with respect to the group $\mathfrak{G}$ associated with $w(z)$, whereas the transforms of the members of class C_1 are never automorphic with respect to $\mathfrak{G}$. Hence no sequence of maps of C_2 can converge to a map of C_1 in this case [3].

There remains but one case to be treated, where $\tilde{F}_w$ is of hyperbolic type. F_w is to be conformally equivalent to the interior of the unit circle punctured at an interior point, which we may assume, without loss of generality, to be the origin. We consider, therefore, the region R defined by: $0 < |w| < 1$. In this case, Theorem 3.1 and the modified proposition for annulus both fail, as the following example shows. Let $W_n(w)$ be defined by

$$W_n(w) = \left(1 - \frac{1}{n}\right)w \qquad (n = 1, 2, \cdots). \tag{3.6}$$

For $0 < |w| < 1$, $W_n(w)$ converges to w. Each $W_n(w)$ is of class C_2 whereas w itself is of class C_1.

Bibliography

1. P. Appell and E. Goursat, *Les Fonctions Algébriques*, vol. II, Paris, 1930.
2. G. Aumann and C. Carathéodory, *Ein Satz über die konforme Abbildung mehrfach zusammenhängender ebener Gebiete*, Mathematische Annalen, vol. 109(1934), pp. 756–763.
3. Maurice H. Heins, *On the iteration of functions analytic and single-valued in a given multiply-connected region*, Bulletin of the American Mathematical Society, vol. 46(1940), p. 44, abstract no. 78.
4. P. Koebe, *Abhandlungen zur Theorie der konformen Abbildung*, IV, Acta Mathematica, vol. 41(1918), pp. 305–344.
5. R. Nevanlinna, *Eindeutige analytische Funktionen*, Berlin, 1936.
6. H. Weyl, *Die Idee der Riemannschen Fläche*, Leipzig, 1923.

The Institute for Advanced Study.

THE ABSOLUTE CONVERGENCE OF TRIGONOMETRICAL SERIES

By Raphaël Salem

The purpose of this paper is to establish some theorems on absolute convergence of trigonometrical series. The first part of the paper is related to some classes of trigonometrical series which cannot converge absolutely at more than one point without being absolutely convergent everywhere. The second part deals with some properties of the sets of points at which a trigonometrical series can converge absolutely without being everywhere absolutely convergent. The third part is devoted to a generalization of the Denjoy-Lusin theorem.

In the first part of the paper, the following theorem is proved:

Theorem I. *If the series*

$$\sum \rho_n \cos (nx - \alpha_n) \qquad (\rho_n \geqq 0) \tag{1}$$

converges absolutely at two points x_0, x_1, *then the series* $\sum \rho_n \mid \sin n(x_1 - x_0) \mid$ *converges.*

This theorem, although very simple, does not seem to have been stated before, and it has some important consequences. It leads to the following theorems:

Theorem II. *The series* (1) *in which the sequence* $\{\rho_n\}$ *is non-increasing cannot converge absolutely at more than one point if* $\sum \rho_n = \infty$.

(Points whose abscissas differ from π are not considered as different.)

Theorem III. *The same theorem is true if instead of supposing the sequence* $\{\rho_n\}$ *non-increasing we suppose that* ρ_{n+p}/ρ_n *is bounded, independently of* n *and* $p > 0$.

Theorem IV. *The series* (1) *in which* $\sum \rho_n = \infty$ *cannot converge absolutely at more than one point if*

$$\sum_1^n \frac{1}{\rho_n} = O(n^2).$$

Theorem V. *If the assumptions on the coefficients* ρ_n *of any one of the Theorems* II, III, *or* IV *are satisfied, the series*

$$\sum \rho_n \cos (k_n x - \alpha_n) \qquad (\sum \rho_n = \infty) \tag{2}$$

Received December 12, 1940 (§§1–12) and January 21, 1941 (§§13–15).

cannot converge absolutely at two points x_0, x_1 whose abscissas differ from δ, provided that the integers $\{k_n\}$ are such that the numbers $k_n\delta$ (mod 2π) are uniformly distributed (gleichverteilt)[1] *on the unit circle.*

It follows in particular from this theorem that Theorems II, III, IV hold for the series

$$\sum \rho_n \cos (n^p x - \alpha_n),$$

p being any integer.

In the second part of the paper we consider the sets of points which we call "of the type N". A set E is of the type N if a trigonometrical series (1) exists which converges absolutely in E, with $\sum \rho_n = \infty$.

Similarly, a set E is of the type N' if a trigonometrical series (1) exists such that $\sum \rho_n \cos^2 (nx - \alpha_n)$ converges in E, with $\sum \rho_n = \infty$.

Plainly every set of the type N is also of the type N'. In answer to a question which has been propounded by Professor Zygmund, we prove the following theorem:

THEOREM VI. *Every set of the type N' is also of the type N.*

We next consider some properties of perfect sets of the type N, and we prove the following theorem:

THEOREM VII. *If a perfect set P is of the type N, then for every function $F(x)$ non-decreasing in $(0, 2\pi)$ constant in every interval contiguous to P but not everywhere*[2] *we have*

$$\overline{\lim} \int_0^{2\pi} \cos 2nx \, dF = F(2\pi) - F(0).$$

This is an extension of a result previously obtained by Professor Zygmund, who has proved that the Fourier-Stieltjes coefficients of dF cannot tend to zero.

We prove then a converse theorem:

THEOREM VIII. *If a perfect set P is such that one function $F(x)$ exists, non-decreasing in $(0, 2\pi)$, constant in every interval contiguous to P, but increasing from one interval to another, and such that*

$$\overline{\lim} \int_0^{2\pi} \cos 2nx \, dF = F(2\pi) - F(0),$$

then a trigonometrical series (1) can be found, with $\sum \rho_n = \infty$, which converges absolutely "almost everywhere" in P, that is to say, in a subset P_1 of P such that the variation of $F(x)$ in $P - P_1$ is zero.

[1] See H. Weyl, *Über die Gleichverteilung von Zahlen mod. Eins*, Math. Ann., vol. 77(1916), pp. 313–352.

[2] The construction of such functions is well known. See E. Hille and J. D. Tamarkin, *Remarks on a known example of a monotone continuous function*, American Mathematical Monthly, vol. 36(1929), pp. 255–264.

Theorems VII and VIII lead to interesting examples of sets which are not of the type N and sets which are "almost everywhere" of the type N.

We next consider the following characteristic of a perfect set P. A positive number η being given, let us denote by q the smallest integer such that P can be covered by q intervals each of length η. Obviously q is non-decreasing when η tends to zero. We prove the following theorem:

THEOREM IX. *If q increases slowly enough when η tends to zero, namely, if $q = o(|\log \eta|)$, then the perfect set P is of the type N.*

In the third part of the paper Theorems X, XI, and XII deal with a generalization of the well-known Denjoy-Lusin theorem on sets of absolute convergence.

I

1. **Proof of Theorem** I. Let x_0, x_1 be two points of absolute convergence of the series (1). From the equality

$$\cos(nx_1 - \alpha_n) = \cos n(x_1 - x_0)\cos(nx_0 - \alpha_n) - \sin n(x_1 - x_0)\sin(nx_0 - \alpha_n)$$

and from the absolute convergence at the points x_0, x_1 we deduce immediately that

$$\sum \rho_n \,|\sin n(x_1 - x_0)\sin(nx_0 - \alpha_n)| < \infty.$$

But from the absolute convergence at x_0 we deduce that

$$\sum \rho_n \,|\sin n(x_1 - x_0)\cos(nx_0 - \alpha_n)| < \infty$$

and by adding the two last inequalities we get

$$\sum \rho_n \,|\sin n(x_1 - x_0)| < \infty. \tag{3}$$

This proves the theorem.

2. **Proof of Theorem** II. This proof is an immediate consequence of the inequality (3) and of a well-known theorem of Fatou[3] which states that if $\sum \rho_n = \infty$ and if the sequence $\{\rho_n\}$ is non-increasing, the series $\sum \rho_n \,|\sin n\delta|$ cannot converge if $\delta \not\equiv 0 \pmod{\pi}$.

Another proof of Theorem II will be included as a particular case in the proof of Theorem V.

3. **Proof of Theorem** III. Let us suppose $\sum \rho_n = \infty$ and $\rho_{n+p}/\rho_n < A$ $(p \geqq 0)$ and let us put

$$r_n = \frac{1}{A}\max(\rho_n, \rho_{n+1}, \cdots);$$

then

$$r_n < \frac{\rho_n}{A}\max\left(\frac{\rho_{n+p}}{\rho_n}\right) < \rho_n.$$

[3] See A. Zygmund, *Trigonometrical Series*, Warsaw, 1935, p. 134.

Hence, if the series (1) converges absolutely at two points x_0, x_1, so does the series $\sum r_n \cos(nx - \alpha_n)$, but plainly $r_{n+1} \leqq r_n$ and as $r_n > A^{-1}\rho_n$, $\sum r_n = \infty$; hence by Theorem II the absolute convergence at x_0, x_1 is impossible if $x_1 - x_0 \not\equiv 0 \pmod{\pi}$.

The same result holds good if instead of supposing ρ_{n+p}/ρ_n bounded, we suppose $\rho_{n+p}/\rho_n < \varphi(n)$, $\varphi(n)$ being an increasing function tending to infinity, such that $\sum \rho_n/\varphi(n) = \infty$. The proof is the same.

4. **Proof of Theorem** IV. If the series (1) converges at x_0 and x_1, then the series (3) converges, and we can find an increasing function $\omega(n)$ tending to infinity such that

$$\sum \rho_n \omega(n) \mid \sin n(x_1 - x_0) \mid < \infty.$$

But[4]

$$\sum_1^n \mid \sin n(x_1 - x_0) \mid = \sum_1^n \frac{1}{(\rho_n \omega(n))^{\frac{1}{2}}} (\rho_n \omega(n))^{\frac{1}{2}} \mid \sin n(x_1 - x_0) \mid$$

$$< \left(\sum_1^n \frac{1}{\rho_n \omega(n)}\right)^{\frac{1}{2}} \left(\sum_1^n \rho_n \omega(n) \sin^2 n(x_1 - x_0)\right)^{\frac{1}{2}}.$$

Hence, if

$$\sum_1^n \frac{1}{\rho_n} = O(n^2),$$

we have

$$\sum_1^n \frac{1}{\rho_n \omega(n)} = o(n^2)$$

and

$$\frac{1}{n} \sum_1^n \mid \sin n(x_1 - x_0) \mid = o(1).$$

This is impossible if $x_1 - x_0 \not\equiv 0 \pmod{\pi}$ because it implies

$$\frac{1}{n} \sum_1^n \sin^2 n(x_1 - x_0) = o(1);$$

that is to say,

$$\lim \frac{1}{n} \sum_1^n \cos 2n(x_1 - x_0) = 1.$$

This is impossible because

$$\left| \sum_1^n \cos 2n(x_1 - x_0) \right| < \frac{1}{\mid \sin (x_1 - x_0) \mid}.$$

Theorem IV is thus proved.

[4] If some of the ρ_n vanish, we can always replace them by the corresponding u_n, $\sum u_n$ being convergent.

5. **Proof of Theorem** V. Let us suppose that the series (2) converges absolutely at two points x_0, x_1 and let us put $x_1 - x_0 = \delta$. By Theorem I, the series $\sum \rho_n \mid \sin k_n\delta \mid$ converges.

(i) Let us suppose that $\{\rho_n\}$ is non-increasing. From the inequality

$$\sum_1^n \rho_n \sin^2 k_n\delta < A,$$

A being a constant, we deduce that

$$\sum_1^n \rho_n \cos 2k_n\delta > \sum_1^n \rho_n - 2A \tag{4}$$

for every n.

But if the numbers $k_n\delta$ (mod 2π) are uniformly distributed on the unit circle, we have, by a well-known theorem of Weyl[5]

$$\frac{1}{n}\sum_1^n \cos 2k_n\delta = o(1),$$

that is to say,

$$\frac{1}{n}\left|\sum_1^n \cos 2k_n\delta\right| < \epsilon_n,$$

ϵ_n tending to zero. Hence, by Abel's transformation

$$\left|\sum_1^n \rho_n \cos 2k_n\delta\right|$$
$$< (\rho_1 - \rho_2)\epsilon_1 + 2(\rho_2 - \rho_3)\epsilon_2 + \cdots + (n-1)(\rho_{n-1} - \rho_n)\epsilon_{n-1} + n\rho_n\epsilon_n,$$

and from this inequality we deduce easily that

$$\sum_1^n \rho_n \cos 2k_n\delta = o(\rho_1 + \cdots + \rho_n).$$

This contradicts (4).

It is interesting to observe that the uniform distribution of the numbers $k_n\delta$ is a condition unnecessarily stringent, and that it is sufficient to suppose

$$\overline{\lim}\,\frac{1}{n}\sum_1^n \cos 2k_n\delta < 1. \tag{5}$$

(ii) If instead of supposing $\{\rho_n\}$ non-increasing we suppose that ρ_{n+p}/ρ_n is bounded ($p > 0$), we reach the same conclusion by the process of Theorem III.

(iii) If we make the assumption $\sum_1^n 1/\rho_n = O(n^2)$, the very proof of Theorem IV shows that the series (2) cannot converge at two points whose abscissas differ from δ if (5) holds good, hence in particular if the numbers $k_n\delta$ are uniformly distributed.

[5] See Weyl, loc. cit.

As an application let us consider the series $\sum \rho_n \cos (n^p x - \alpha_n)$, p being any integer not less than 1. It is known that the numbers $n^p \delta$ are uniformly distributed (mod 2π) if δ/π is irrational. If δ/π is rational, it is easily seen that

$$\overline{\lim} \frac{1}{n} \sum_1^n \cos 2n^p \delta < 1.$$

Hence in the hypothesis of any one of Theorems II, III, IV, the series $\sum \rho_n \cos (n^p x - \alpha_n)$ cannot converge absolutely at more than one point.

Theorem V is thus proved.

II

We next establish some properties of the sets of type N. Theorem I is useful in this study because it shows that sets of absolute convergence of the series (1) can be obtained by a translation of sets of absolute convergence of the series $\sum \rho_n \sin nx$ which is of a simpler type.

6. **Proof of Theorem** VI. Let E be a set of the type N'. Then there exists a series $\sum \rho_n \cos (k_n x - \alpha_n)$ such that, in E,

$$\sum \rho_n \cos^2 (k_n x - \alpha_n) < \infty, \qquad \sum \rho_n = \infty, \qquad \rho_n > 0.$$

Let us put

$$S_n = \sum_1^n \rho_n \quad \text{and} \quad u_n = S_n^{\frac{1}{2}} - S_{n-1}^{\frac{1}{2}}.$$

Plainly $\sum u_n$ diverges. We have

$$\sum_1^n u_n \,|\cos (k_n x - \alpha_n)| = \sum_1^n \frac{u_n}{\rho_n^{\frac{1}{2}}} \rho_n^{\frac{1}{2}} \,|\cos (k_n x - \alpha_n)|$$

$$< \left(\sum_1^n \frac{u_n^2}{\rho_n}\right)^{\frac{1}{2}} \left(\sum_1^n \rho_n \cos^2 (k_n x - \alpha_n)\right)^{\frac{1}{2}}.$$

Now

$$\frac{u_n^2}{\rho_n} = \frac{(S_n^{\frac{1}{2}} - S_{n-1}^{\frac{1}{2}})^2}{S_n - S_{n-1}} = \frac{S_n^{\frac{1}{2}} - S_{n-1}^{\frac{1}{2}}}{S_n^{\frac{1}{2}} + S_n^{\frac{1}{2}} S_{n-1}^{\frac{1}{2}} + S_{n-1}^{\frac{1}{2}}} < \frac{1}{S_{n-1}^{\frac{1}{2}}} - \frac{1}{S_n^{\frac{1}{2}}}.$$

Hence $\sum u_n^2/\rho_n$ converges, and this proves the convergence of $\sum u_n \,|\cos (k_n x - \alpha_n)|$ in E. Hence E is of the type N, and this proves the theorem.

7. **Proof of Theorem** VII. Let P be a perfect set of the type N. If x_0 is a point of P and P_0 denotes the set P translated by $-x_0$, then if the series (1) converges absolutely in P, the series $\sum \rho_n \,|\sin nx|$ converges for every x belonging to P_0 and so does the series $\sum \rho_n \sin^2 nx$. Hence the difference

$$\sum_1^n \rho_n - \sum_1^n \rho_n \cos 2nx$$

is bounded (not uniformly) for every x belonging to P_0, and putting

$$R_n(x) = \frac{\rho_1 \cos 2x + \rho_2 \cos 4x + \cdots + \rho_n \cos 2nx}{\rho_1 + \rho_2 + \cdots + \rho_n},$$

we have, for every x belonging to P_0, owing to the divergence of $\sum \rho_n$,

$$\lim R_n(x) = 1 \tag{6}$$

(not uniformly). The quantities $R_n(x)$ being uniformly bounded ($| R_n(x) | \leqq 1$), we can take the Stieltjes integral of both sides of (6) with respect to a function F, non-decreasing in $(0, 2\pi)$, constant in every interval contiguous to P_0 but not everywhere. We thus get

$$\lim_{n\to\infty} \frac{\rho_1 \int_0^{2\pi} \cos 2x\, dF + \rho_2 \int_0^{2\pi} \cos 4x\, dF + \cdots + \rho_n \int_0^{2\pi} \cos 2nx\, dF}{\rho_1 + \rho_2 + \cdots + \rho_n} = F(2\pi) - F(0).$$

This is impossible unless

$$\overline{\lim} \int_0^{2\pi} \cos 2nx\, dF = F(2\pi) - F(0), \tag{7}$$

and this proves the theorem. We can observe, integrating by parts, that this condition is equivalent to

$$\underline{\lim} \left| n \int_0^{2\pi} F(x) \sin 2nx\, dx \right| = 0.$$

8. **Proof of Theorem** VIII. Let P be a perfect set in $(0, 2\pi)$ and F a function of the type defined in the proof of Theorem VII, constant in every interval contiguous to P, but increasing from one interval to another, and such that the equality (7) holds good. We have

$$\overline{\lim} \int_0^{2\pi} (1 - 2 \sin^2 nx)\, dF = F(2\pi) - F(0),$$

that is to say,

$$\underline{\lim} \int_0^{2\pi} \sin^2 nx\, dF = 0.$$

Hence there is a sequence $\{k_n\}$ of integers such that

$$\int_0^{2\pi} \sin^2 k_n x\, dF < \epsilon_n,$$

ϵ_n decreasing and tending to zero. We can always find a positive sequence $\{\rho_n\}$ such that $\sum \rho_n = \infty$, $\sum \rho_n \epsilon_n < \infty$. We can take, for example, $\rho_n = \epsilon_n^{-\frac{1}{2}} - \epsilon_{n-1}^{-\frac{1}{2}}$. We have then

$$\sum_1^\infty \int_0^{2\pi} \rho_n \sin^2 k_n x\, dF < \infty;$$

the integrated quantities being all positive, we conclude that the series

$$\sum_1^\infty \rho_n \sin^2 k_n x$$

converges in a subset P_1 of P such that the variation of F over $P - P_1$ is zero. By Theorem VI this is also true for a series $\sum u_n \mid \sin k_n x \mid$ with $\sum u_n = \infty$, and this proves Theorem VIII.

9. To give an application of Theorems VII and VIII we shall consider the perfect set constructed in $(0, 2\pi)$ in the following manner.

We first divide the interval in three parts of lengths proportional to ξ_1, $1 - 2\xi_1$, ξ_1 respectively, and we remove the central part.

Each one of the intervals left is divided in three parts of lengths proportional to ξ_2, $1 - 2\xi_2$, ξ_2, and the central parts are removed.

In the p-th operation each of the 2^{p-1} remaining intervals are divided in three parts of lengths proportional to ξ_p, $1 - 2\xi_p$, ξ_p, and the 2^{p-1} central parts are removed.

We continue these operations indefinitely, the sequence $\{\xi_p\}$ being such that $0 < \xi_p \leqq \frac{1}{2}$. It is plain that we obtain thus a perfect set P nowhere dense. It is easily seen that the points of P are represented by the infinite series

$$x = 2\pi(\theta_1 + \xi_1\theta_2 + \xi_1\xi_2\theta_3 + \cdots + \xi_1\xi_2 \cdots \xi_{p-1}\theta_p + \cdots), \tag{8}$$

where θ_i is equal to zero, or to $1 - \xi_i$.

The 2^{p-1} intervals removed in the p-th operation will be denoted by δ_{pk} ($k = 1, 2, \cdots, 2^{p-1}$); each of them is of length $2\pi\xi_1 \cdots \xi_{p-1}(1 - 2\xi_p)$. The right-hand end-points of the intervals δ_{pk} are obtained by making in (8) $\theta_p = 1 - \xi_p$ and $\theta_i = 0$ for each $i > p$. The left-hand end-points are obtained by making $\theta_p = \xi_p$ and $\theta_i = 0$ for each $i > p$, but we can preserve the form (8) by making $\theta_p = 0$ and $\theta_i = 1 - \xi_i$ for each $i > p$.

After p operations $2^p - 1$ intervals are removed and 2^p intervals are left which we denote by η_{pk} ($k = 1, 2, \cdots, 2^p$); each of them is of length $2\pi\xi_1\xi_2 \cdots \xi_p$. Thus the measure of P is

$$2\pi \lim 2^p \xi_1\xi_2 \cdots \xi_p,$$

and we suppose that this limit is zero.

Now we define a continuous function $F(x)$ as being equal, for x belonging to P and given by (8), to

$$\frac{a_1}{2} + \frac{a_2}{2^2} + \cdots + \frac{a_p}{2^p} + \cdots,$$

where

$$a_i = 0 \quad \text{if} \quad \theta_i = 0 \qquad \text{and} \qquad a_i = 1 \quad \text{if} \quad \theta_i = 1 - \xi_i.$$

There is no change to make in the well-known argument about Cantor's ternary set to show that $F(x)$ is a continuous non-decreasing function[6] constant in every interval contiguous to P but increasing from one interval to another.

Following exactly the method of Hille and Tamarkin we can calculate the Stieltjes integral

$$I = \int_0^{2\pi} e^{nix}\, dF(x).$$

Let us consider the 2^p left-hand end-points α_{pk} of the 2^p intervals η_{pk} ($k = 1, 2, \cdots, 2^p$). Their abscissas are obtained by making $\theta_i = 0$ or $\theta_i = 1 - \xi_i$ ($i = 1, 2, \cdots, p$) in the *finite* expression

$$\alpha_{pk} = 2\pi(\theta_1 + \xi_1\theta_2 + \cdots + \xi_1\xi_2 \cdots \xi_{p-1}\theta_p).$$

Let us divide $(0, 2\pi)$ in intervals of length tending to zero, each of them containing either one complete interval η_{pk} or no part of them. Then, as in the intervals δ_{jk} ($j = 1, 2, \cdots, p$), F is constant, and since in each η_{pk} it increases by 2^{-p}, we have for approximate value of I

$$I_p = \frac{1}{2^p} \sum \exp \{2\pi ni[\theta_1 + \xi_1\theta_2 + \cdots + \xi_1\xi_2 \cdots \xi_{p-1}\theta_p]\},$$

the summation being extended to all combinations of $\theta_i = 0$ and $\theta_i = 1 - \xi_i$. We have

$$\begin{aligned} I_p &= \frac{1}{2^p} \prod_{k=1}^{p} \{1 + \exp [2\pi ni\xi_1\xi_2 \cdots \xi_{k-1}(1 - \xi_k)]\} \\ &= \prod_{k=1}^{p} \exp [\pi ni\xi_1\xi_2 \cdots \xi_{k-1}(1 - \xi_k)] \cos \pi n\xi_1\xi_2 \cdots \xi_{k-1}(1 - \xi_k), \end{aligned}$$

but $1 - \xi_1 + \xi_1(1 - \xi_2) + \xi_1\xi_2(1 - \xi_3) + \cdots = 1$. Hence

$$I = e^{\pi ni} \prod_{k=1}^{\infty} \cos \pi n\xi_1\xi_2 \cdots \xi_{k-1}(1 - \xi_k).$$

Theorem VII shows that the set P cannot be of the type N unless

$$\overline{\lim_{n\to\infty}} \prod_{k=1}^{\infty} \cos 2\pi n\xi_1\xi_2 \cdots \xi_{k-1}(1 - \xi_k) = 1. \tag{9}$$

It is seen immediately, for example, that P cannot be of the type N if all the ξ_i are equal to ξ.[7] For n being given we can always find a positive integer k such that

$$\mu \leqq 2n\xi^{k-1}(1 - \xi) \leqq 1 - \mu,$$

[6] See Hille and Tamarkin, loc. cit. We have adopted the notations of this paper.

[7] This result is known. See V. Niemytzki, *Sur quelques classes d'ensembles linéaires avec applications aux séries trigonométriques absolument convergentes*, Rec. Soc. Math. Moscou, vol. 33(1926), pp. 5–32.

if μ is small enough. It is sufficient to have

$$\frac{\log(\mu^{-1}-1)}{|\log \xi|} \geqq 1,$$

that is to say,

$$\mu \leqq \frac{1}{1+\xi^{-1}},$$

and μ being thus chosen, we shall have

$$\left| \prod_{k=1}^{\infty} \cos 2\pi n \xi_1 \xi_2 \cdots \xi_{k-1}(1-\xi_k) \right| < \cos \pi\mu.$$

We shall show now that if

$$(\xi_1 \xi_2 \cdots \xi_k)^{1/k} = o(k^{-\frac{1}{2}}),$$

the relation (9) holds; and then, by Theorem VIII, the set P is "almost everywhere" of the type N.

Let us suppose, in fact, that

$$(\xi_1 \xi_2 \cdots \xi_k)^{1/k} < \frac{1}{\varphi(k)k^{\frac{1}{2}}},$$

$\varphi(k)$ being an increasing function, tending to infinity as slowly as we please. Putting $\xi_1\xi_2 \cdots \xi_{k-1}(1-\xi_k) = \zeta_k$, we have

$$\prod_{k=1}^{\infty} \cos 2\pi n\zeta_k = \prod_{k=1}^{p} (1 - 2\sin^2 \pi n\zeta_k) \prod_{k=p+1}^{\infty} (1 - 2\sin^2 \pi n \zeta_k).$$

By Dirichlet's theorem we can find an integer n such that $A \leqq n \leqq At^p$ and such that $|\sin \pi n\zeta_k| < t^{-1}$ for $k = 1, 2, \cdots, p$. Let us take $A = p$, $t = [p\varphi(p)]^{\frac{1}{2}}$. We have

$$p \leqq n \leqq p[p\varphi(p)]^{\frac{1}{2}p} \tag{10}$$

and

$$\prod_{k=1}^{p} (1 - 2\sin^2 \pi n\zeta_k) > \left(1 - \frac{2}{p\varphi(p)}\right)^p. \tag{11}$$

On the other hand as $1 - u > e^{-2u}$ if $0 < u < \frac{1}{2}$ we have

$$\prod_{p+1}^{\infty} (1 - 2\sin^2 \pi n\zeta_k) > \exp\left[-4\pi^2 n^2(\zeta_{p+1}^2 + \zeta_{p+2}^2 + \cdots)\right]$$

(provided $4\pi^2 n^2 \zeta_{k+1}^2 < 1$ for $k \geqq p$ which is true, as it will be seen in a moment). But

$$\zeta_k^2 = \xi_1^2 \cdots \xi_{k-1}^2(1-\xi_k)^2 < \xi_1^2 \cdots \xi_{k-1}^2(1-\xi_k^2),$$

hence

$$\sum_{p+1}^{\infty} \zeta_k^2 < \xi_1^2 \cdots \xi_p^2$$

and

$$n^2 \xi_1^2 \cdots \xi_p^2 < p^2 [p\varphi(p)]^p \left[\frac{1}{p\varphi^2(p)}\right]^p = \frac{p^2}{[\varphi(p)]^p}.$$

Then

$$\prod_{p+1}^{\infty} (1 - 2 \sin^2 \pi n \zeta_k) > \exp\left\{-4\pi^2 \frac{p^2}{[\varphi(p)]^p}\right\}. \tag{12}$$

If p is allowed to increase infinitely, the inequalities (11) and (12) together with (10) prove that

$$\overline{\lim_{n=\infty}} \prod_{k=1}^{\infty} \cos 2\pi n \zeta_k = 1.$$

This is the result as stated.

10. The condition (7) which has been proved necessary in order that a perfect set P should be of the type N can be put in another form which is closer to the structure of the set itself.

Let us construct the set P by the following process: We remove first the greatest contiguous interval (if there are several intervals of the same greatest length, we take the first on the left); we get thus two intervals containing P; from each of them we remove the greatest contiguous interval contained, and so on. After p operations we have removed $1 + 2 + \cdots + 2^{p-1} = 2^p - 1$ contiguous intervals and there are 2^p intervals left, which cover P and which we denote by η_{pk} $(k = 1, 2, \cdots, 2^p)$.

We define a continuous function $F_p(x)$ by the conditions $F_p(0) = 0$, $F_p(2\pi) = 1$; $F_p = 2^{-p}k$ in the k-th contiguous interval counted from the left to the right existing after p operations $(k = 1, 2, \cdots, 2^p - 1)$, $F_p(x)$ linear between two such successive contiguous intervals. It is easy to see that $F_p(x)$ tends to a non-decreasing continuous function $F(x)$ constant in every interval contiguous to P. Now

$$\int_0^{2\pi} \sin^2 nx \, dF = \sum_k \int_{\eta_{pk}} \sin^2 nx \, dF.$$

But if α_{pk} is the left-hand end-point of η_{pk},

$$\left| \int_{\eta_{pk}} \sin^2 nx \, dF - \sin^2 n\alpha_{pk} \int_{\eta_{pk}} dF \right| < 2n\eta_{pk} \int_{\eta_{pk}} dF = 2n\eta_{pk} \frac{1}{2^p}.$$

Hence

$$\left| \int_0^{2\pi} \sin^2 nx \, dF - \frac{1}{2^p} \sum_k \sin^2 n\alpha_{pk} \right| < 2n \frac{1}{2^p} \sum_k \eta_{pk}.$$

Now, if (7) holds, there exists an increasing sequence of integers $\{n_m\}$ such that $\int_0^{2\pi} \sin^2 n_m x\, dF < \epsilon_m \to 0$. To every n_m corresponds a p_m such that for every $p \geqq p_m$ we have $n_m 2^{-p} \sum_k \eta_{pk} < \epsilon_m$, hence also $2^{-p} \sum_k \sin^2 n_m \alpha_{pk} < 3\epsilon_m$.

This means that with every p increasing infinitely we can associate an integer n_p increasing infinitely such that

$$\begin{cases} n_p \dfrac{1}{2^p} \sum_k \eta_{pk} = o(1), \\ \dfrac{1}{2^p} \sum_k \sin^2 n_p \alpha_{pk} = o(1). \end{cases}$$

By Schwarz' inequality we can replace $\sin^2 n_p\alpha_{pk}$ by $|\sin n_p\alpha_{pk}|$ hence by $\{n_p\alpha_{pk}\pi^{-1}\}$, $\{z\}$ denoting the positive distance between z and the nearest integer. If we denote by η_p the mean-value of the 2^p intervals η_{pk}, we see that if P is of the type N there exists a sequence of integers n_p such that

$$\begin{cases} n_p \eta_p = o(1), \\ \dfrac{1}{2^p} \sum_k \left\{ n_p \dfrac{\alpha_{pk}}{\pi} \right\} = o(1). \end{cases}$$

These conditions give a rough idea of the relation existing between the arithmetical properties of the numbers α_{pk} and the "thickness" of a set of the type N. If the α_{pk}/π can be "easily" approximated in mean by rational numbers (i.e., by rationals with a denominator increasing slowly with p), then n_p will increase slowly and η_p can decrease slowly; that is to say, the set can be comparatively "thick". If, on the contrary, the α_{pk}/π do not lend themselves to easy approximation, n_p can increase as rapidly as its upper bound shown by Dirichlet's theorem, and, accordingly, η_p must decrease rapidly.

11. We shall now define a class of perfect sets which are of the type N.

Proof of Theorem IX. With every perfect set we can associate a function $q(\eta)$, q being the smallest integer such that P can be covered with q intervals each of length η. $q(\eta)$ is non-decreasing when η tends to zero.

Let α_i $(i = 1, 2, \cdots, q)$ be the abscissas of the left-hand end-points of the q intervals of length η covering P. ϵ_q being positive and tending to zero as $q \to \infty$, we can always, by Dirichlet's theorem, find an integer m_q such that

$$\frac{1}{\epsilon_q} \leqq m_q \leqq \frac{1}{\epsilon_q}\left(\frac{1}{\epsilon_q}\right)^q \tag{13}$$

and such that

$$|\sin m_q\alpha_i| < \pi\epsilon_q \qquad (i = 1, 2, \cdots, q).$$

If m_q is such that $m_q\eta < \epsilon_q$, we shall have

$$|\sin m_q x| < (\pi + 1)\epsilon_q$$

for every x belonging to P. Since by (13) the sequence of m_q is unbounded, we can extract an increasing sequence of integers n_q such that $|\sin n_q x| < \epsilon_q'$ for every x belonging to P. Hence, by a suitable choice of ρ_q as shown in §8, the series $\sum \rho_q \sin n_q x$ will converge absolutely in P with $\sum \rho_q = \infty$.

It remains to find a sufficient condition in order that m_q satisfying (13) be such that $m_q \eta < \epsilon_q$. It is sufficient to have

$$\eta < \epsilon_q^{q+2}$$

and, since ϵ_q can tend to zero as slowly as we please, it is sufficient to have

$$\eta^{1/q} = o(1),$$

a relation which is satisfied if $q = o(|\log \eta|)$. Theorem IX is thus proved.

12. The sufficient condition which has just been found has some interest because there exist sets which are *not* of the type N and for which

$$q = o\left(\frac{1}{\eta}\right)^{\alpha},$$

α being a positive number as small as we please.

This is seen immediately in considering the sets studied in §9 and in which all the ξ_i are equal to the same number ξ.

For those sets, which are not of the type N; we have for $\eta = 2\pi\xi^k$, $q = 2^k$; hence also $q \leqq 2^k$ for $2\pi\xi^k < \eta < 2\pi\xi^{k-1}$. From this we deduce easily that

$$q < C\left(\frac{1}{\eta}\right)^{\log 2/|\log \xi|},$$

C being a constant. Since we can take ξ as near to zero as we please, the result follows.

It will be interesting to show now that there are sets for which

$$q > \left(\frac{1}{\eta}\right)^{\beta},$$

β being less than 1 but as near 1 as we please, and which are of the type N.

Let us take the set P considered in §9 and let us put $\xi_p = \frac{1}{2}$ for every p except for a sequence $\{i_j\}$ which we shall define in a moment and for which $\xi_{i_j} = 1/2j$. Let us take $p = i_j$ and let us cover the set with 2^p intervals, each of length $\eta_p = 2\pi\xi_1 \cdots \xi_p = 2\pi 2^{-p}(j!)^{-1}$. Now if α_{pk} are the origins of the 2^p intervals η_{pk} ($k = 1, 2, \cdots, 2^p$), it is seen immediately that if we take

$$n_p = 2^p(j-1)!$$

we have

$$\sin n_p \alpha_{pk} = O\left(\frac{1}{j}\right) \qquad (k = 1, 2, \cdots, 2^p)$$

and

$$n_p \eta_p = O\left(\frac{1}{j}\right).$$

Hence

$$\sin n_p x = O\left(\frac{1}{j}\right)$$

for every x belonging to P. Hence, as it has been proved in §11, the set P is of the type N.

But for this set if we take

$$2\pi\xi_1\xi_2 \cdots \xi_{i_{j+1}} < \eta < 2\pi\xi_1\xi_2 \cdots \xi_{i_j},$$

it is easy to see that

$$q > 2^{i_j},$$

but

$$\frac{1}{\eta} < \frac{1}{\xi_1\xi_2 \cdots \xi_{i_{j+1}}} = 2^{i_{j+1}}(j+1)!.$$

Hence, to have $q > (1/\eta)^\beta$ it is sufficient to choose the sequence i_j such that

$$2^{i_j} > [2^{i_{j+1}}(j+1)!]^\beta.$$

It will be sufficient to have

$$i_j \log 2 > \beta i_{j+1} \log 2 + \beta(j+1) \log (j+1)$$

or

$$i_j\left[1 - \beta\,\frac{i_{j+1}}{i_j}\right] > \frac{1}{\log 2}\,(j+1) \log (j+1).$$

This inequality will hold, for instance, if we take $i_j = j^2$; and $\beta < 1$ can be as near to 1 as we please. This proves the result as stated.

III

13. Let us consider the trigonometrical series

$$(S) \qquad \sum_1^\infty \rho_n \cos (nx - \alpha_n) \qquad (\rho_n \geqq 0).$$

We shall henceforth suppose, in this paper, that $\sum \rho_n = \infty$. The Denjoy-Lusin theorem states that the set of points for which

$$\sum_1^\infty \rho_n \mid \cos (nx - \alpha_n) \mid < \infty$$

is of measure zero.

We shall prove the following theorem.

THEOREM X. *The set of points E for which*

$$\underline{\lim} \frac{\sum_1^n \rho_n \,|\cos(nx - \alpha_n)|}{\sum_1^n \rho_n} \leqq \alpha \tag{14}$$

is of measure zero if $\alpha < 2/\pi$.

Let $f(x)$ be the characteristic function of E. The functions

$$\frac{\sum_1^n \rho_n \,|\cos(nx - \alpha_n)|}{\sum_1^n \rho_n} \tag{15}$$

being uniformly bounded, we have,[8] if (14) holds,

$$\underline{\lim} \frac{\sum_1^n \rho_n \int_0^{2\pi} f(x)\,|\cos(nx - \alpha_n)|\,dx}{\sum_1^n \rho_n} \leqq \alpha \int_0^{2\pi} f(x)\,dx. \tag{16}$$

Now

$$\int_0^{2\pi} f(x)\,|\cos(nx - \alpha_n)|\,dx = \int_0^{2\pi} f\left(x + \frac{\alpha_n}{n}\right)|\cos nx|\,dx;$$

if we observe that

$$\int_0^{2\pi} \left| f\left(x + \frac{\alpha_n}{n}\right) - f(x) \right| dx$$

tends to zero for $n = \infty$, and that $\int_0^{2\pi} f(x)\,|\cos nx|\,dx$ tends[9] to

$$\frac{1}{2\pi} \int_0^{2\pi} f(x)\,dx \int_0^{2\pi} |\cos x|\,dx,$$

we have

$$\lim_{n=\infty} \int_0^{2\pi} f(x)\,|\cos(nx - \alpha_n)|\,dx = \frac{2}{\pi} \int_0^{2\pi} f(x)\,dx.$$

Hence

$$\lim_{n=\infty} \frac{\sum_1^n \rho_n \int_0^{2\pi} f(x)\,|\cos(nx - \alpha_n)|\,dx}{\sum_1^n \rho_n} = \frac{2}{\pi} \int_0^{2\pi} f(x)\,dx. \tag{17}$$

[8] See, e.g., E. W. Hobson, *The Theory of Functions of a Real Variable*, 2d ed., vol. II, p. 318.

[9] See A. Zygmund, op. cit., p. 173.

Comparing (16) and (17), we see that, if $\alpha < 2/\pi$, we have $\int_0^{2\pi} f(x)\,dx = 0$, and this proves the theorem.

14. The set in which (14) holds contains obviously the set of absolute convergence of the series (S). Hence, it can contain a perfect subset. We shall prove the following theorem:

THEOREM XI. *If* (14) *holds in a perfect set P, then every bounded function F non-decreasing in* $(0, 2\pi)$ *constant in each interval contiguous to P but not everywhere is such that* (*even if F is continuous*)

$$\overline{\lim} \left| \int_0^{2\pi} e^{nix}\,dF \right| > 0.$$

Let us assume that a function F of the above described type is such that the Fourier-Stieltjes coefficients of dF tend to zero. We have

$$|\cos(nx - \alpha_n)| = \sum_{k=-\infty}^{+\infty} d_k e^{iknx} \qquad \left(d_0 = \frac{2}{\pi}\right)$$

with $|d_k| < k^{-2}A$, A being a constant independent of k and n. Taking into account that $\left| \int_0^{2\pi} e^{pix}\,dF \right|$ is bounded and tends to zero for $p = \infty$, we have immediately

$$\lim_{n=\infty} \int_0^{2\pi} |\cos(nx - \alpha_n)|\,dF = \frac{2}{\pi}[F(2\pi) - F(0)]. \tag{18}$$

But if (14) holds in P, we have

$$\overline{\lim} \frac{\sum_1^n \rho_n \int_0^{2\pi} |\cos(nx - \alpha_n)|\,dF}{\sum_1^n \rho_n} \leqq \alpha[F(2\pi) - F(0)], \tag{19}$$

the functions (15) being uniformly bounded; and (18) contradicts (19) if $\alpha < 2/\pi$. Hence the theorem is proved.

15. The constant $2/\pi$ in Theorem X is the best possible one. If $\alpha = 2/\pi$, the theorem is no longer true. This will result from the following theorem:

THEOREM XII. *If $\rho_n = O(1)$, we have, almost everywhere,*

$$\lim \frac{\sum_1^n \rho_n |\cos(nx - \alpha_n)|}{\sum_1^n \rho_n} = \frac{2}{\pi}.$$

We shall first prove the following lemma:

LEMMA. *If the* γ_n *are complex quantities such that* $\sum_1^\infty |\gamma_n| = \infty$, *and* $|\gamma_n| = O(1)$, *we have, almost everywhere,*

$$R_n(x) = \frac{\sum_1^n \gamma_n e^{inx}}{\sum_1^n |\gamma_n|} \to 0.$$

The proof of this lemma will be the same as Weyl's proof of uniform distribution of the numbers $n_k x$ for almost all x, $\{n_k\}$ being an increasing sequence of integers.[10]

We have

$$\int_0^{2\pi} |R_n(x)|^2\,dx = 2\pi \frac{\sum_1^n |\gamma_n|^2}{(\sum_1^n |\gamma_n|)^2} < \frac{M}{\sum_1^n |\gamma_n|}$$

if $2\pi |\gamma_n| < M$. Let us take a sequence of integers $\{n_k\}$ such that

$$k^2 \leqq \sum_1^{n_k} |\gamma_n| < (k+1)^2.$$

Then $\sum \int_0^{2\pi} |R_{n_k}(x)|^2\,dx$ converges, and consequently $\sum |R_{n_k}(x)|$ converges almost everywhere, and thus

$$R_{n_k}(x) \to 0$$

almost everywhere. Now let us take an n such that

$$n_k \leqq n < n_{k+1}.$$

We have

$$\left| R_n(x) \sum_1^n |\gamma_n| - R_{n_k}(x) \sum_1^{n_k} |\gamma_n| \right| < \sum_{n_k+1}^{n_{k+1}} |\gamma_n|.$$

Hence

$$\left| R_n(x) - \frac{\sum_1^{n_k} |\gamma_n|}{\sum_1^n |\gamma_n|} R_{n_k}(x) \right| \leqq \frac{\sum_{n_k+1}^{n_{k+1}} |\gamma_n|}{\sum_1^{n_k} |\gamma_n|} = \frac{\sum_1^{n_{k+1}} |\gamma_n|}{\sum_1^{n_k} |\gamma_n|} - 1 < \frac{(k+2)^2}{k^2} - 1,$$

and this proves that $R_n(x) \to 0$ almost everywhere.

[10] See Weyl, loc. cit.

Remark. The assumption $|\gamma_n| = O(1)$ could be replaced by a more general one; but it is necessary to make some hypothesis about the coefficients γ_n. It is easily seen, for instance, that

$$\frac{1}{2^n}\sum_1^n 2^k e^{kix}$$

does not tend to zero almost everywhere for $n = \infty$.

We can now prove Theorem XII. We have

$$|\cos x| = \sum_{k=-\infty}^{+\infty} c_k e^{ikx} \qquad \left(c_0 = \frac{2}{\pi}\right)$$

with $|c_k| = O(k^{-2})$. Hence

$$|\cos(nx - \alpha_n)| = \sum_{k=-\infty}^{+\infty} c_k e^{-ik\alpha_n} e^{iknx}$$

and

$$\frac{\sum_1^n \rho_n |\cos(nx - \alpha_n)|}{\sum_1^n \rho_n} = \sum_{k=-\infty}^{+\infty} c_k Q_{k,n}(x),$$

where

$$Q_{k,n}(x) = \frac{\sum_1^n \rho_n e^{-ik\alpha_n} \cdot e^{iknx}}{\sum_1^n \rho_n}.$$

Applying the lemma, we see that k being given $\neq 0$, if x does not belong to a certain set E_k of measure zero, we have

$$\lim_{n=\infty} Q_{k,n}(x) = 0. \tag{20}$$

Hence, there is a set $\mathfrak{S}$ of measure 2π such that for every x belonging to $\mathfrak{S}$, and for every integer $k \neq 0$, (20) holds good. Taking into account that the $|Q_{k,n}(x)|$ are uniformly bounded, and that $|c_k| = O(k^{-2})$, we see immediately that

$$\frac{\sum_1^n \rho_n |\cos(nx - \alpha_n)|}{\sum_1^n \rho_n} \to c_0 = \frac{2}{\pi}$$

for every x belonging to $\mathfrak{S}$. This proves the theorem.

MONTREAL, CANADA.

THE DISTRIBUTION OF THE NUMBER OF SUMMANDS IN THE PARTITIONS OF A POSITIVE INTEGER

By Paul Erdös and Joseph Lehner

1. It is well known that $p(n)$, the number of unrestricted partitions of a positive integer n, is given by the asymptotic formula [2][1]

$$p(n) \sim \frac{1}{4n3^{\frac{1}{2}}} \exp Cn^{\frac{1}{2}}, \qquad C = \pi(\tfrac{2}{3})^{\frac{1}{2}}. \tag{1.1}$$

In §2 we prove that the "normal" number of summands in the partitions of n is $C^{-1}n^{\frac{1}{2}} \log n$. More precisely, we prove the following

Theorem 1.1. *Denote by $p_k(n)$ the number of partitions of n which have at most k summands. Then, for*

$$k = C^{-1}n^{\frac{1}{2}} \log n + xn^{\frac{1}{2}}, \tag{1.2}$$

we have

$$\lim_{n\to\infty} \frac{p_k(n)}{p(n)} = \exp\left(-\frac{2}{C}e^{-\frac{1}{2}Cx}\right). \tag{1.3}$$

The right member of (1.3) is strictly monotone and continuous; it tends to 0 as $x \to -\infty$ and to 1 as $x \to +\infty$. Hence, it is a distribution function. Also from (1.3) we clearly obtain the weaker result that if $f(n)$ is any function tending with n to infinity, then the number of summands in "almost all" partitions of n lies between

$$\frac{n^{\frac{1}{2}} \log n}{C} \pm f(n) \cdot n^{\frac{1}{2}}. \tag{1.4}$$

It is easily seen that the number of partitions of n having k or less summands is equal to the number of partitions of n in which no summand exceeds k. Thus the preceding results can be applied to this case also.

In §3 we consider $P(n)$, the number of partitions of n into unequal parts. (By a theorem of Euler, $P(n)$ is also equal to the number of partitions of n into odd summands with repetitions allowed.) We obtain results similar to the above for $p_k(n)$, but we shall not give all details of the proof.

In §4 we derive an asymptotic formula for $p_k(n)$,

$$p_k(n) \sim \frac{\binom{n-1}{k-1}}{k!}, \tag{1.5}$$

valid uniformly in k in the range $k = o(n^{\frac{1}{3}})$.

Received January 9, 1941. The first-mentioned author is a Harrison Research Fellow.

[1] Numbers in brackets refer to the bibliography at the end of this paper.

These matters, to our knowledge, have not been discussed previously. Somewhat similar questions have been suggested by Castelnuovo [1] and treated by Tricomi [5]. The collected works of Sylvester are full of papers dealing with $p_k(n)$, for particular values of k. However, Sylvester did not consider the effect of making k a function of n, i.e., he did not discuss the asymptotic behavior of $p_k(n)$. His attack was entirely algebraic. In their famous paper on partitions Hardy and Ramanujan [2] give an inequality for $p_k(n)$ for finite k. If we use the generating function for $p_k(n)$ and the calculus of residues, it is easy to derive an asymptotic formula (see §5).

In one of his numerous papers on partitions Sylvester ([4], pp. 90–99, esp. p. 93, footnote) remarked that in attempting to work out problems of this sort one meets with another class of partitions in the midst of the problem, so that it is difficult to avoid circularity. It has been possible to do this in our case by using elementary inequalities for the occurring partition function.

2. We start from the following identity

$$
\begin{aligned}
p_k(n) &= p(n) - \sum_{1 \leqq r \leqq n-k} p(n-(k+r)) \\
&\quad + \sum_{\substack{0<r_1<r_2 \\ 1<r_1+r_2 \leqq n-2k}} p(n-(k+r_1)-(k+r_2)) \\
&\quad - \sum_{\substack{0<r_1<r_2<r_3 \\ 1<r_1+r_2+r_3 \leqq n-3k}} p(n-(k+r_1)-(k+r_2)-(k+r_3)) + - \cdots \\
&= p(n)\{1 - S_1 + S_2 - S_3 + - \cdots\}.
\end{aligned} \tag{2.1}
$$

(2.1) is a simple application of the Sieve of Eratosthenes; we use also the remark in the paragraph of §1 following (1.4), and the obvious fact[2] that the number of partitions of n into summands which include k is equal to $p(n-k)$. Also, by a well-known principle of Bruns' method ([3], p. 75, (59)),

$$
1 - S_1 + S_2 - + \cdots - S_{2\nu-1} \leqq \frac{p_k(n)}{p(n)} \leqq 1 - S_1 + S_2 - + \cdots + S_{2\nu} \qquad (\nu = 1, 2, 3, \cdots). \tag{2.2}
$$

Now we estimate $S_1, S_2, \cdots$. Using (1.1), we have, with $k = C^{-1}n^{\frac{1}{2}} \log n + xn^{\frac{1}{2}}$,

$$
S_1 \sim \sum_{1 \leqq r \leqq n-k-1} \frac{n}{n-k-r} \exp\left[C(n-k-r)^{\frac{1}{2}} - Cn^{\frac{1}{2}}\right] = \sum_{r \leqq n^{\frac{3}{4}}}{}_1 + \sum_{r > n^{\frac{3}{4}}}{}_2 .
$$

In $\sum_1$, $n(n-k-r)^{-1} \sim 1$ and $(n-k-1)^{\frac{1}{2}} \sim n^{\frac{1}{2}} - \frac{1}{2}n^{-\frac{1}{2}}(k+r)$; thus

[2] This principle will be used several times in this paper.

$$\textstyle\sum_1 \sim \sum_{r \leq n^{3/5}} \exp\,[-C \cdot \frac{1}{2} n^{-1/2}(k + r)] = n^{-1/2} \exp\,[-\frac{1}{2}Cx] \sum_{1 \leq r \leq n^{3/5}} \exp\,[-\frac{1}{2}Crn^{-1/2}]$$

$$= n^{-1/2} \exp\,[-\tfrac{1}{2}Cx] \exp\,[-\tfrac{1}{2}Cn^{-1/2}] \frac{1 - \exp\,[-\frac{1}{2}Cn^{1/10}]}{1 - \exp\,[-\frac{1}{2}Cn^{-1/2}]}$$

$$\sim n^{-1/2} \exp\,[-\tfrac{1}{2}Cx] \cdot \frac{2n^{1/2}}{C};$$

$$\textstyle\sum_2 < n \sum_{r > n^{3/5}} \exp\,[-\frac{1}{2}Cn^{-1/2}(k + r)] < n \sum_{r > n^{3/5}} \exp\,[-\frac{1}{2}Crn^{-1/2}]$$

$$\textstyle < n \exp\,[-\frac{1}{2}Cn^{-1/2}n^{3/5}] \sum_{r > n^{3/5}} 1 < n^2 \exp\,[-\frac{1}{2}Cn^{1/10}] = o(1).$$

Therefore,

$$S_1 \sim \frac{2}{C} \exp\,[-\tfrac{1}{2}Cx]. \tag{2.3}$$

Next

$$S_2 = \frac{1}{2!\,p(n)} \sum_{1 \leq r_1, r_2 \leq n-2k} p(n - 2k - r_1 - r_2) - \frac{1}{p(n)} \sum_{1 \leq r \leq n-2k} p(n - 2k - 2r)$$

$$\textstyle = \frac{1}{2!}\left(\sum_1 + \sum_2\right) - \sum_{3\; r \leq n^{3/5}} - \sum_{4\; r > n^{3/5}},$$

where $\sum_1$ runs over all pairs (r_1, r_2) in which neither r_1 nor r_2 exceeds $n^{3/5}$; $\sum_2$ over all pairs in which at least one member exceeds $n^{3/5}$. As before, we find

$$\textstyle\sum_1 \sim \frac{1}{n} \exp\,[-Cx] \sum_{r_1, r_2 \leq n^{3/5}} \exp\,[-\frac{1}{2}Cn^{-1/2}(r_1 + r_2)]$$

$$\textstyle = \frac{1}{n} \exp\,[-Cx]\left(\sum_{r_1 \leq n^{3/5}} \exp\,[-\frac{1}{2}Cr_1 n^{-1/2}]\right)^2 \sim \left(\frac{2}{C} \exp\,[-\frac{1}{2}Cx]\right)^2,$$

$$\textstyle\sum_2 = o(1),$$

$$\textstyle\sum_3 = \frac{1}{n} \exp\,[-Cx] \sum_{r \leq n^{3/5}} \exp\,[-Crn^{-1/2}] \sim C^{-1} n^{-1/2} \exp\,[-Cx]$$

$$= o(1),$$

$$\textstyle\sum_4 = o(1).$$

Therefore,

$$S_2 \sim \frac{1}{2!}\left(\frac{2}{C} \exp\,[-\tfrac{1}{2}Cx]\right)^2. \tag{2.4}$$

Similarly we get for S_ν

$$S_\nu \sim \frac{1}{\nu!}\left(\frac{2}{C} \exp\,[-\tfrac{1}{2}Cx]\right)^\nu. \tag{2.5}$$

Hence from (2.2) and the fact that $S_\nu \to 0$ with ν^{-1}, we have

$$\frac{p_k(n)}{p(n)} \sim 1 + \sum_{\nu=1}^{\infty} (-1)^\nu S_\nu = \exp\left(-\frac{2}{C} e^{-\frac{1}{2}cx}\right),$$

which is (1.3).

3. We now consider $P(n)$, the number of partitions of n into unequal summands. Such a partition will be called an "unequal partition"; a partition into odd summands we shall call an "odd partition". We outline the proof of the following

Theorem 3.1. *For almost all unequal partitions of n, the number of summands in a given partition not exceeding $xn^{\frac{1}{2}}$ lies between*

$$\frac{2n^{\frac{1}{2}}}{D} \log \frac{2}{1 + e^{-Dx}} \pm \epsilon n^{\frac{1}{2}}, \qquad D = \pi(\tfrac{1}{3})^{\frac{1}{2}}. \tag{3.1}$$

To the odd partition

$$n = 1 \cdot x_1 + 3 \cdot x_3 + \cdots + (2r + 1)x_{2r+1} \tag{3.21}$$

corresponds in a one-to-one way the unequal partition

$$n = \sum_{l=0}^{r} (2l + 1) \sum_{t=1}^{s_{2l+1}} 2^{\alpha_{2l+1,t}} = \sum_{l,t} (2l + 1)2^{\alpha_{2l+1,t}}, \tag{3.22}$$

where[3]

$$x_i = 2^{\alpha_{i,1}} + 2^{\alpha_{i,2}} + \cdots + 2^{\alpha_{i,s_i}}.$$

Denote by $A(x)$ the number of summands not exceeding $xn^{\frac{1}{2}}$ in a given partition of n, and by $\sum_{P(n)}$ a sum which runs over all unequal partitions of n. Then

$$\sum_{P(n)} A(x) = \sum_{1 \leq u < xn^{\frac{1}{2}}} P_u(n), \tag{3.3}$$

where $P_u(n)$ is the number of unequal partitions of n which contain the summand u. Let $u = 2^k(2v + 1)$.

In order to calculate $P_u(n)$, we consider all odd partitions of n (3.21) which contain $(2v + 1)$,

$$n = 1 \cdot x_1 + \cdots + (2v + 1)x_{2v+1} + \cdots, \tag{3.41}$$

and in which, moreover, 2^k occurs in the dyadic expansion of x_{2v+1},

$$x_{2v+1} = \cdots + 2^k + \cdots. \tag{3.42}$$

By the correspondence (3.21), (3.22), $P_u(n)$ is equal to the number of such partitions.

In order to count these partitions we let $k = 0, 1, 2, \cdots$ in turn. $k = 0$

[3] This correspondence, of course, furnishes a proof of Euler's theorem.

implies u is odd. Then in (3.41), x_{2v+1} runs through all odd integers, since in (3.42) a $2^0 = 1$ must occur. Hence, we are interested in those odd partitions which contain $2v + 1$ exactly once, exactly three times, etc. Their number is clearly

$$P(n - (2v + 1)) - P(n - 2(2v + 1)) + P(n - 3(2v + 1)) - P(n - 4(2v + 1)) + \cdots, \tag{3.51}$$

and this must be summed on $v = 0, 1, 2, \cdots$, such that $u = 2v + 1 < xn^{\frac{1}{2}}$.

In the same way we count, for a general k, those odd partitions which contain $2v + 1$ exactly $2^k, 2^k + 1, \cdots, 2^{k+1} - 1$ times; $2^{k+1} + 2^k, 2^{k+1} + 2^k + 1, \cdots, 2^{k+2}$ times; etc. The number of such partitions is seen to be

$$P(n - 2^k(2v + 1)) - P(n - 2\cdot 2^k(2v + 1)) + P(n - 3\cdot 2^k(2v + 1)) - P(n - 4\cdot 2^k(2v + 1)) + \cdots, \tag{3.52}$$

this to be summed on $v = 0, 1, \cdots$, such that $u = 2^k(2v + 1) < xn^{\frac{1}{2}}$.

To these sums we can apply the method of §2, using the asymptotic expression for $P(n)$ given by Hardy-Ramanujan ([2], p. 113),[4]

$$P(n) \sim \frac{\exp\,[Dn^{\frac{1}{2}}]}{4\cdot 3^{\frac{1}{4}}\cdot n^{\frac{3}{4}}}. \tag{3.61}$$

In this way we obtain the asymptotic value of $\sum_{P(n)} A(x)$ as

$$\sum_{P(n)} A(x) \sim \frac{2n^{\frac{1}{2}}}{D} P(n) \log \frac{2}{1 + e^{-Dx}}. \tag{3.62}$$

Next we consider

$$\begin{aligned} \Delta(x) &= \sum_{P(n)} [A(x) - n^{\frac{1}{2}}F(x)]^2 \\ &\sim \sum_{P(n)} A^2(x) - nP(n)F^2(x), \end{aligned} \tag{3.71}$$

where we have written for abbreviation

$$F(x) = \frac{2}{D} \log \frac{2}{1 + e^{-Dx}}.$$

Now

$$\sum_{P(n)} A^2(x) = \sum_{1 \leq u_1, u_2 < xn^{\frac{1}{2}}} P_{u_1, u_2}(n), \tag{3.72}$$

where $P_{u_1, u_2}(n)$ denotes the number of unequal partitions of n containing both

[4] See also L. K. Hua, *On the number of partitions of a number into unequal parts*, Bulletin of the American Mathematical Society, vol. 46(1940), p. 419, abstract no. 279.

u_1 and u_2 ; $P_{u_1,u_1}(n) = P_{u_1}(n)$. We calculate $P_{u_1,u_2}(n)$ by the same methods used to find $P_u(n)$. It turns out that[5]

$$(3.73) \qquad P_{u_1,u_2}(n) \sim E_1 \cdot E_2 \cdot P(n), \qquad E_1 = \frac{P_{u_1}(n)}{P(n)}, \qquad E_2 = \frac{P_{u_2}(n)}{P(n)};$$

thus

$$(3.74) \qquad \sum_{P(n)} A^2(x) \sim nF^2(x)P(n)$$

and

$$(3.75) \qquad \Delta(x) = o(nF^2(x)P(n)).$$

For a fixed $\epsilon > 0$, let $N(x, \epsilon)$ be the number of unequal partitions of n for which

$$| A(x) - n^{\frac{1}{2}}F(x) | > \epsilon n^{\frac{1}{2}}.$$

Then

$$\Delta(x) > N(x, \epsilon) \cdot \epsilon^2 n,$$

and by (3.75),

$$N(x, \epsilon) = o(P(n)).$$

This is equivalent to Theorem 3.1.

This leads to the following $(x \to \infty)$

THEOREM 3.2. *For almost all unequal partitions of n the number of summands in a given partition lies between*

$$\frac{2n^{\frac{1}{2}}}{D} \log 2 \pm \epsilon n^{\frac{1}{2}}.$$

By sharper arguments we can obtain

THEOREM 3.3. *The number of unequal partitions of n in which the number of summands in a given partition is less than*

$$\frac{2n^{\frac{1}{2}}}{D} \log 2 + yn^{\frac{1}{4}}$$

is given by a Gaussian integral.

We add the following two theorems, which may be of some interest. They can be proved very easily by using the methods of this section.

THEOREM 3.4. *Let*

$$(3.81) \qquad n = a_1 + a_2 + \cdots + a_k$$

[5] (3.73) expresses the independence, in the sense of probability, of the function $P_u(n)/(P(n))^2$. This holds, however, only for the values considered, i.e., $u_1, u_2 < xn^{\frac{1}{2}}$.

be any partition of n. Define

$$f(n; a_1, a_2, \cdots, a_k) = f(n) = \sum_i A_i,$$

where A_i runs over the different summands in the given partition. Then for almost all partitions $f(n)$ lies between

$$\frac{6n}{\pi^2}(1 \pm \epsilon). \tag{3.82}$$

THEOREM 3.5. *Let $\varphi(n; a_1, a_2, \cdots, a_k) = \varphi(n)$ denote the number of different summands in the partition* (3.81). *Then for almost all partitions $\varphi(n)$ lies between*

$$\frac{2n^{1/2}}{C}(1 \pm \epsilon), \qquad C = \pi(\tfrac{2}{3})^{1/2}. \tag{3.9}$$

4. We now discuss the asymptotic behavior of $p_k(n)$ for $k = o(n^{1/3})$ and prove the following

THEOREM 4.1.

$$p_k(n) \sim \frac{\binom{n-1}{k-1}}{k!}, \tag{4.1}$$

this formula being valid uniformly in k for $k = o(n^{1/3})$.[6]

LEMMA 4.2. *Let $k = o(n^{1/3})$. Then*[7]

$$p_k(n) > \frac{1}{2} \cdot \frac{n}{k^2} \cdot p_{k-1}(n). \tag{4.2}$$

In the proof of this lemma, we shall consider partitions into exactly k summands some of which may be zero. This is equivalent to the case of partitions into k or fewer summands.

First we show that

$$p_k(n) > \frac{n}{k^3} p_{k-1}(n). \tag{4.31}$$

Let

$$n = a_1 + a_2 + \cdots + a_{k-1}, \qquad 0 \leqq a_1 \leqq a_2 \leqq \cdots \leqq a_{k-1}, \tag{4.32}$$

be any partition of n into $k - 1$ parts. Clearly $a_{k-1} > n/k$. Now if we write $a_{k-1} = x + y$, $0 \leqq x \leqq y$, we obtain from each partition (4.32) at least

[6] I.e., for every $\epsilon > 0$, and $0 < k^3 n^{-1} < \epsilon$, the ratio of $p_k(n)$ to $\binom{n-1}{k-1}/k!$ remains between $1 \pm \epsilon$ as $n \to \infty$.

[7] This result no doubt holds for $k = O(n^{1/2})$.

$a_{k-1}/2 > n/2k$ partitions of n into k parts. Hence, from all partitions (4.32) we get at least $p_{k-1}(n) \cdot n/2k$ partitions of n into k parts,

$$n = b_1 + b_2 + \cdots + b_k, \qquad 0 \leqq b_1 \leqq b_2 \leqq \cdots \leqq b_k. \tag{4.33}$$

In the set (4.33) no partition is duplicated more than $\binom{k}{2}$ times; therefore

$$p_{k-1}(n) \cdot \frac{n}{2k} \leqq p_k(n) \cdot \binom{k}{2},$$

and (4.31) follows.

Next, in (4.32), let $A_1, A_2, \cdots, A_r$ be the distinct positive summands, $0 < A_1 < \cdots < A_r$. If we break up each A_i into two parts as in the preceding paragraph, we obtain at least

$$\tfrac{1}{2}(A_1 + A_2 + \cdots + A_r) \tag{4.41}$$

partitions in (4.33).

In the following we denote by $\sum_{p_k(n)}$ a sum which runs over all partitions of n into k parts some of which may be zero. We shall estimate $\sum_{p_{k-1}(n)} \sum_{i=1}^{r} A_i$.

We have

$$\sum_{p_{k-1}(n)} \sum_{i=1}^{r} A_i = \sum_{s=1}^{n} s p_{k-2}(n - s), \tag{4.42}$$

since a given integer s appears in the left member as many times as there are partitions of n into $k - 1$ parts one of which is s, i.e., just $p_{k-2}(n - s)$ times. By an extension of the same reasoning we get

$$\begin{aligned} \sum_{p_{k-1}(n)} \sum_{i=1}^{k-1} a_i &= n p_{k-1}(n) \\ &= \sum_{s=1}^{n} s\{p_{k-2}(n - s) + p_{k-3}(n - 2s) + p_{k-4}(n - 3s) + \cdots\}, \end{aligned} \tag{4.43}$$

the series in the braces terminating of its own accord. Now

$$p_{k-3}(n - 2s) + p_{k-4}(n - 3s) + \cdots < 3p_{k-2}(n - s). \tag{4.44}$$

For, clearly,

$$\begin{aligned} &p_{k-3}(n - 2s) \leqq p_{k-2}(n - s), \\ &p_{k-4}(n - 3s) \leqq p_{k-3}(n - s), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &p_{k-u}(n - (u - 1)s) \leqq p_{k-u+1}(n - s) \end{aligned}$$

(see footnote 2); hence

$$p_{k-3}(n - 2s) + p_{k-4}(n - 3s) + \cdots \leqq p_{k-2}(n - s) + p_{k-3}(n - s) + \cdots.$$

Applying (4.31) to the last inequality, we see that the left member does not exceed

$$p_{k-2}(n-s)\left\{1+\frac{(k-2)^3}{n-s}+\frac{(k-3)^3(k-2)^3}{(n-s)^2}+\cdots\right\}.$$

We remark that we need only consider $s < \frac{1}{2}n$, for otherwise the right member of (4.43) reduces to the first term. For $s < \frac{1}{2}n$, the above expression in braces is less than

$$1+\frac{2k^3}{n}+\left(\frac{2k^3}{n}\right)^2+\cdots < 1+\tfrac{1}{2}+\tfrac{1}{4}+\cdots = 2,$$

since $k = o(n^{1/3})$. This proves (4.44).

Finally, (4.42), (4.43), (4.44) give

$$\begin{aligned}\sum_{p_{k-1}(n)}\sum_{i=1}^{r} A_i &= \sum_{s=1}^{n} s p_{k-2}(n-s)\\ &> \frac{1}{4}\sum_{s=1}^{n} s\{p_{k-2}(n-s)+p_{k-3}(n-2s)+\cdots\} \qquad (4.45)\\ &= \tfrac{1}{4} n p_{k-1}(n).\end{aligned}$$

(4.41) and (4.45) mean that by the process of breaking up each A_i into two parts we obtain from the set (4.32) at least $\frac{1}{8} n p_{k-1}(n)$ partitions in (4.33). Moreover, no partition is duplicated more than $\binom{k}{2}$ times. Hence

$$\tfrac{1}{4} n p_{k-1}(n) < p_k(n)\cdot\binom{k}{2},$$

and this proves Lemma 4.2.

COROLLARY 4.3. *If $k = o(n^{1/3})$, then the number of partitions of n into exactly k positive summands is asymptotically equal to the number of partitions of n into k or fewer positive summands.*

For by a t-fold application of Lemma 4.2, we have

$$p_{k-t}(n) < \left(\frac{2k^2}{n}\right)^t p_k(n);$$

hence

$$\sum_{t=1}^{k-1} p_{k-t}(n) < p_k(n)\sum_{t=1}^{k-1}\left(\frac{2k^2}{n}\right)^t = o(p_k(n)),$$

since $k = o(n^{1/3})$.

LEMMA 4.4. *The number of partitions of n into exactly k positive summands not all of which are different is $o(p_k(n))$.*

Let any such partition be given by

$$n = t_1b_1 + t_2b_2 + \cdots + t_xb_x, \qquad \sum_{i=1}^{x} t_i = k, \tag{4.51}$$

and $t_i > 1$ for some i, i.e., $x < k$. To this we make correspond

$$n = c_1 + c_2 + \cdots + c_x, \qquad c_i = t_ib_i. \tag{4.52}$$

This furnishes a single-valued mapping of (4.51) into a subset of the set of partitions of n with fewer than k summands. This inverse mapping is far from being single-valued, however. In fact, given a fixed partition of (4.52),

$$n = d_1 + d_2 + \cdots + d_{k-t}, \qquad t > 0, \tag{4.53}$$

we inquire in how many ways it can be mapped into (4.51). The inverse mapping exhausts the set (4.51).

For this purpose we select v of the d's, say $d_{i_1}, d_{i_2}, \cdots, d_{i_v}$, and split d_{i_1} into w_1 equal parts, d_{i_2} into w_2 equal parts, $\cdots$, d_{i_v} into w_v equal parts ($w_1 \geqq 2, \cdots, w_v \geqq 2$).[8] We must evidently have

$$w_1 + w_2 + \cdots + w_v = v + t. \tag{4.54}$$

Since in a given decomposition $v \leqq t$, we get

$$w_1 + w_2 + \cdots + w_v \leqq 2t. \tag{4.55}$$

Hence, the total number of decompositions obtainable from all possible choices of v and $w_1, w_2, \cdots, w_v$ is less than[9]

$$p(1) + p(2) + \cdots + p(2t) < 4^t. \tag{4.56}$$

From a given decomposition (4.54) we obtain at most $(k - t)^v \leqq (k - t)^t < k^t$ partitions in (4.51), so that, all in all, we get at most 4^tk^t partitions in (4.51) from our fixed partition (4.53). But for each t there are $p_{k-t}(n)$ partitions of the form (4.53); hence the total number of partitions of n into k positive summands not all of which are different is less than

$$\sum_{t=1}^{k-1} 4^t k^t p_{k-t}(n),$$

and by Lemma 4.2 this is less than

$$p_k(n) \sum_{t=1}^{k-1} 4^t k^t \left(\frac{2k^2}{n}\right)^t = p_k(n) \sum_{t=1}^{k-1} \left(\frac{8k^3}{n}\right)^t = o(p_k(n)),$$

by virtue of the condition $k = o(n^{\frac{1}{3}})$. Thus Lemma 4.4 is proved.

[8] We assume here $w_1 \mid d_{i_1}, \cdots, w_v \mid d_{i_v}$. This assumption only strengthens the inequalities which follow.

[9] This estimate follows from an elementary inequality for $p(n)$, $p(n) < 2^{n-1}$. For the proof of the latter, see footnote 11.

Lemma 4.5.[10] *The number, $p'_k(n)$, of partitions of n into k positive summands whose order is considered (i.e., two partitions are counted as different if they differ only in the order of their summands) is* $\binom{n-1}{k-1}$.

Let

$$n = a_1 + a_2 + \cdots + a_k\,, \qquad a_i > 0. \tag{4.61}$$

To this *partition* we make correspond the *combination*

$$a_1\,, a_1 + a_2\,, a_1 + a_2 + a_3\,, \cdots, a_1 + a_2 + \cdots + a_{k-1}\,, \tag{4.62}$$

and this correspondence is clearly one to one. But each of the $k - 1$ integers in (4.62) is not greater than $n - 1$, since $a_k \geqq 1$.[11]

Now we can prove Theorem 4.1. From Corollary 4.3, it is clear that we need consider only partitions having exactly k positive summands. Moreover, from Lemma 4.4, we see that we may assume all summands in a given partition to be different. But from a partition in which all k summands are different we obtain $k!$ partitions of the type considered in Lemma 4.5. Thus the theorem follows.

5. By the application of the Hardy-Littlewood method we can obtain a second proof of Theorem 4.1. But it hardly seems worth while to use this elaborate method unless something more results. It is easily seen that the essential contribution is furnished by the neighborhood of $x = 1$. Hence what we need is information about the asymptotic character of the generating function

$$\frac{1}{(1-x)(1-x^2)\cdots(1-x^k)} = 1 + \sum_1^\infty p_k(n)x^n$$

around $x = 1$. The possibility of obtaining a suitably sharp asymptotic representation remains to be investigated.

Bibliography

1. G. Castelnuovo, *Sur quelques problèmes se rattachant au calcul des probabilités*, Annales de l'Institut Henri Poincaré, vol. 3(1932–33), pp. 465–490.
2. G. H. Hardy and S. Ramanujan, *Asymptotic formulae in combinatory analysis*, Proceedings of the London Mathematical Society, vol. 17(1918), pp. 75–115.
3. E. Landau, *Vorlesungen über Zahlentheorie*, vol. I, Leipzig, 1927.
4. J. J. Sylvester, *On the Partition of Numbers*, Collected Works, vol. II, Cambridge, 1908.
5. F. Tricomi, *Su di una variabile casuale connessa con un notevole tipo di partizioni di un numero intero*, Giornale Istituto Ital. Attuari, vol. 2(1931), pp. 455–468.

University of Pennsylvania.

[10] This lemma and proof are well known.

[11] The estimate for $p(n)$ given in footnote 9 follows easily from Lemma 4.5. For

$$p(n) < p'_1(n) + p'_2(n) + \cdots + p'_n(n) = 2^{n-1}.$$

RIESZ SUMMABILITY METHODS OF ORDER r, FOR $\mathcal{R}(r) < 0$

By G. E. Forsythe

Given a series $\sum_{n=0}^{\infty} u_n$ of complex terms and a complex parameter r, whose real part will be denoted by $\mathcal{R}(r)$, let

$$(1) \qquad \alpha_n = \sum_{k=0}^{n-1} \left(1 - \frac{k}{n}\right)^r u_k \qquad (n = 1, 2, \cdots),$$

and let

$$\beta(t) = \sum_{k=0}^{[t]-1} \left(1 - \frac{k}{t}\right)^r u_k \qquad (1 \leqq t < \infty).$$

By a^r, where $a > 0$, will always be meant $\exp [r \log a]$, where $\log a$ is given its real value. If $\lim_{n\to\infty} \alpha_n = L$, then $\sum u_n$ is said to be summable-A_r to L. If $\lim_{t\to\infty} \beta(t) = L$, then $\sum u_n$ is said to be summable-B_r to L. These summability methods are due to M. Riesz, and this notation is due to Agnew.[1]

If $-1 < r \leqq 1$, then A_r, B_r and the Cesàro method C_r are all equivalent,[2] while for some values of $r > 1$, A_r and B_r are not equivalent.[3] For $\mathcal{R}(r) < -1$, A_r and B_r are not equivalent.[4] For other values of r, the question of the equivalence of A_r and B_r seems not to be discussed in the literature.

The object of this note is to give a criterion for the equivalence of A_r and B_r for $\mathcal{R}(r) < 0$, based on Agnew's work, and to apply this criterion to show that A_{-1+ih} and B_{-1+ih} $(-\infty < h < \infty)$ are equivalent if and only if $h = 0$. It follows[5] that A_{-1} can take the position $r = -1$ in the scale of Cesàro summability methods C_r.

Let $\varphi_r(x) = \sum_{n=1}^{\infty} n^r x^n$, for $|x| < 1$. $\varphi_r(x)$ has a simple zero at the origin, so that we can define the coefficients $\{e_n^{(r)}\}$ by

$$(2) \qquad f_r(x) = \frac{1}{\varphi_r(x)} - \frac{1}{x} = \sum_{n=0}^{\infty} e_n^{(r)} x^n, \qquad \text{for } |x| < R,\ R > 0.$$

Let $e_{-1}^{(r)} = 1$, for all r.

Received January 17, 1941; presented to the American Mathematical Society, October 26, 1940. The author is a Sigma Xi Fellow from Swarthmore College.

[1] R. P. Agnew, *On Riesz and Cesàro methods of summability*, Transactions of the American Mathematical Society, vol. 35(1933), pp. 532–548. (See this paper for references to Riesz.)

[2] Agnew, op. cit., p. 544; and M. Riesz, *Sur l'équivalence de certaines méthodes de sommation*, Proceedings of the London Mathematical Society, (2), vol. 22(1923–24), pp. 412–419; p. 418.

[3] Agnew, loc. cit., Theorem 4.4, and Riesz, loc. cit., p. 418.

[4] Agnew, loc. cit., Theorem 10.2.

[5] Agnew, loc. cit., p. 544.

THEOREM 1. *If* $\mathcal{R}(r) < 0$, *the methods* A_r *and* B_r *are equivalent if and only if the constants* $\{e_n^{(r)}\}$ *of* (2) *satisfy the following two conditions*:

$$\lim_{n\to\infty} n^{-r} e_n^{(r)} = 0; \tag{3}$$

$$\sum_{k=1}^{n+1} \left| \left(\frac{k}{n}\right)^r \right| \cdot | e_{n-k}^{(r)} | \leqq M \qquad (n = 1, 2, \cdots). \tag{4}$$

Given any $\{\alpha_n\}$, we can always solve (1) uniquely for $\{u_n\}$. To do this, we let

$$u(x) = \sum_{n=0}^{\infty} u_n x^n, \qquad \text{formally.} \tag{5}$$

From (1) we have $n^r \alpha_n = \sum_{k=0}^{n-1} (n - k)^r u_k$. Then, formally,

$$\sum_{n=1}^{\infty} \alpha_n n^r x^n = \sum_{n=1}^{\infty} x^n \sum_{k=0}^{n-1} (n - k)^r u_k = u(x)\varphi_r(x).$$

Then

$$u(x) = \{u(x)\varphi_r(x)\} \cdot \{\varphi_r(x)\}^{-1} = \Big\{ \sum_{n=1}^{\infty} n^r \alpha_n x^n \Big\} \cdot \Big\{ \sum_{n=-1}^{+\infty} e_n^{(r)} x^n \Big\},$$

formally; or

$$u(x) = \sum_{n=0}^{\infty} x^n \sum_{k=1}^{n+1} k^r e_{n-k}^{(r)} \alpha_k . \tag{6}$$

Equating coefficients in (5) and (6), we obtain $u_n = \sum_{k=1}^{n+1} k^r e_{n-k}^{(r)} \alpha_k$ as our solution, so that

$$n^{-r} u_n = \sum_{k=1}^{n+1} \left(\frac{k}{n}\right)^r e_{n-k}^{(r)} \alpha_k = \sum_{k=1}^{\infty} d_{nk} \alpha_k \qquad (n = 1, 2, \cdots), \tag{7}$$

where the matrix $\| d_{nk} \|$ is defined by (7).

Now suppose that $\alpha_n \to L$. If we set $u_0' = u_0 - L$, and $u_n' = u_n$ $(n \geqq 1)$, we get the corresponding $\alpha_n' = \alpha_n - L$ $(n \geqq 1)$ from (1). Hence we may suppose that $L = 0$ without loss of generality in determining u_n for $n \geqq 1$.

From a theorem of Agnew[6] it follows immediately that A_r and B_r are equivalent for $\mathcal{R}(r) < 0$, if and only if $\lim_{n\to\infty} \alpha_n = L$ implies that $\lim_{n\to\infty} n^{-r} u_n = 0$. As remarked we may assume $L = 0$, and by (7) we see that A_r and B_r are equivalent if and only if $\| d_{nk} \|$ is a null-preserving matrix.[7] But for this it is

[6] Agnew, loc. cit., p. 536, Theorem 5.1.

[7] A matrix $\| c_{nk} \|$ is called null-preserving if $\lim_{k\to\infty} x_k = 0$ implies that $\lim_{n\to\infty} \sum_{k=1}^{\infty} c_{nk} x_k = 0$.

necessary and sufficient[8] that we have both

$$\lim_{n\to\infty} d_{nk} = k^r \lim_{n\to\infty} n^{-r} e^{(r)}_{n-k} = 0 \qquad (k = 1, 2, \cdots) \tag{8}$$

and

$$\sum_{k=1}^{\infty} | d_{nk} | \leqq M \qquad (n = 1, 2, \cdots). \tag{9}$$

But (8) is obviously equivalent to (3), while (9) is identical to (4), and Theorem 1 is proved.

THEOREM 2. *For* $-\infty < h < \infty$ *the methods* A_{-1+ih} *and* B_{-1+ih} *are equivalent if and only if* $h = 0$.

I. $h = 0$. Then $\varphi_{-1}(x) = -\log(1 - x)$, and by (2) $\sum_{n=0}^{\infty} e_n^{(-1)} x^n = -1/x - 1/\log(1 - x)$. These constants $\{e_n^{(-1)}\}$ are well known. We have[9]

$$e_n^{(-1)} = (-1)^{n+1} \int_0^1 \frac{x(x-1)\cdots(x-n)}{(n+1)!}\, dx = O\left(\frac{1}{n \log^2 n}\right), \tag{10}$$

whence

$$ne_n^{(-1)} \to 0, \qquad \text{as } n \to \infty. \tag{11}$$

From (1), if $u_0 = 1$ and $u_n = 0$ $(n \geqq 1)$, we have $\alpha_n = 1$ $(n \geqq 1)$ for all r. Hence from (7) we have

$$\sum_{k=1}^{n+1} \left(\frac{k}{n}\right)^r e^{(r)}_{n-k} = 0 \qquad (n = 1, 2, \cdots; \text{all } r). \tag{12}$$

$e_{-1}^{(-1)} = 1$, while from (10) we have $e_n^{(-1)} < 0$ for $n \geqq 0$. Hence

$$\begin{aligned} \sum_{k=1}^{n+1} \left| \left(\frac{k}{n}\right)^{-1} \right| \cdot | e^{(-1)}_{n-k} | &= -\sum_{k=1}^{n} \left(\frac{k}{n}\right)^{-1} e^{(-1)}_{n-k} + \frac{n}{n+1} \\ &= \frac{2n}{n+1} - \sum_{k=1}^{n+1} \left(\frac{k}{n}\right)^{-1} e^{(-1)}_{n-k} \\ &= \frac{2n}{n+1}, \qquad \text{from (12) for } r = -1, \\ &\leqq 2 \qquad (n = 1, 2, \cdots). \end{aligned} \tag{13}$$

But (11) and (13) are (3) and (4) for $r = -1$. Hence A_{-1} and B_{-1} are equivalent, by Theorem 1.

II. $h \neq 0$. The analytic function $\varphi_r(x)$, whose principal branch is defined

[8] T. Kojima, *On generalized Toeplitz's theorems on limit and their applications*, Tôhoku Mathematical Journal, vol. 12(1917), pp. 291–326; p. 300.

[9] J. D. Tamarkin, *Problem 3276*, American Mathematical Monthly, vol. 35(1928), p. 500.

for $|x| < 1$ by $\sum_{n=1}^{\infty} n^r x^n$, has singularities only at 0, 1 and ∞, and the principal branch is regular at 0.[10] For $r \neq -1, -2, \cdots$ and for $|\log x| < 2\pi$ we have[10]

$$\varphi_r(x) = \Gamma(1 + r)\left(\log \frac{1}{x}\right)^{-r-1} + \sum_{\nu=0}^{\infty} \frac{\zeta(-r-\nu)(\log x)^\nu}{\nu!},$$

whence for $r = -1 + ih$ we get for the principal branch

$$\varphi_{-1+ih}(x) - \Gamma(ih)\left(\log \frac{1}{x}\right)^{-ih} \to \zeta(1 - ih), \qquad \text{as } x \to 1 - 0, \tag{14}$$

so that $\varphi_{-1+ih}(x)$ oscillates finitely as $x \to 1 - 0$, without approaching a limit. Fix $r = -1 + ih$.

Now if $\varphi_r(x)$ has a zero x_0 with $0 < |x_0| < 1$, then the series (2) has a radius of convergence $R \leqq |x_0| < 1$. From the Cauchy-Hadamard theorem[11] it follows readily that $\overline{\lim_n} |ne_n^{(r)}| = \infty$, whence (3) fails, so that A_r and B_r are not equivalent in this case.

If $\varphi_r(x)$ has no zero for $0 < |x| < 1$, then $R = 1$ in (2), since 1 is an essential singularity of $\varphi_r(x)$. Suppose, if possible, that $\sum_{n=0}^{\infty} e_n^{(r)} = A$. Then by Abel's theorem[12] we would have $f_r(x) \to A$ as $x \to 1 - 0$, by (2). But this cannot happen, by (14) and the definition of $f_r(x)$.

Hence $\sum_{n=0}^{\infty} e_n^{(r)}$ diverges, and so $\sum_{n=0}^{\infty} |e_n^{(r)}| = \infty$. Therefore

$$\sum_{k=1}^{n+1} \left|\left(\frac{k}{n}\right)^r\right| \cdot |e_{n-k}^{(r)}| > \sum_{k=1}^{n+1} |e_{n-k}^{(r)}| \frac{n}{k} > \sum_{k=0}^{n-1} |e_k^{(r)}| \to \infty, \qquad \text{as } n \to \infty.$$

So (4) fails, and A_r and B_r are not equivalent in this case. This proves Theorem 2 completely.

The question of the equivalence of A_r and B_r for $-1 < \Re(r)$, r not real, is still open. Although Theorem 1 is a possible tool when $-1 < \Re(r) < 0$, the problem of determining the zeros of $\varphi_r(x)$ makes the question appear difficult to answer.

BROWN UNIVERSITY.

[10] E. L. Lindelöf, *Le Calcul des Résidus*, Paris, 1905, pp. 138 ff.
[11] K. Knopp, *Funktionentheorie* I, Sammlung Göschen, Berlin, 1937, p. 68.
[12] E. C. Titchmarsh, *The Theory of Functions*, Oxford, 1932, p. 9.

UNSTABLE MINIMAL SURFACES OF HIGHER TOPOLOGICAL STRUCTURE

By Marston Morse and C. Tompkins

1. **Introduction.** We are concerned with extending the calculus of variations in the large to multiple integrals. The problem of the existence of minimal surfaces of unstable type contains many of the typical difficulties, especially those of a topological nature. Having studied this problem for the case of one boundary [11], [12], we turned to the case of m boundaries. We discovered new difficulties not found either in the general theory when $m = 1$ or in the extensive minimum theory when $m > 1$. The case $m = 2$, however, appeared to contain the essentially new difficulties, and in order to present the relevant new ideas in their simplest form we have kept to this case.

We shall illustrate our results by a theorem which might have been conjectured by Newton. Let g_0 and g_1 be two parallel circles with planes orthogonal to their line of centers. Two such circles sufficiently near together bound a minimal surface of revolution of minimum area. This surface is generated by a segment of a catenary, and is always accompanied by another minimal surface of revolution not of minimum type. This classical result admits a simple generalization.

First recall that a simple, closed, rectifiable curve g is said to satisfy the *chord arc condition* 1a if the ratio of the length of an arbitrary chord of g to the minimum of the corresponding arc lengths of g is bounded from zero. A surface S is said to be a *disc surface* (*ring surface*) if S is given as the continuous image of a disc (circular ring). The above theorems on minimal surfaces of revolution admit the following generalization.

Let g_0 and g_1 be simple, rectifiable, closed curves in n-space satisfying the chord arc condition, separated by an $(n - 1)$-plane and possessing convex projections on suitably chosen $(n - 1)$-planes. If g_0 and g_1 bound a ring minimal surface belonging to a minimizing set, g_0 and g_1 also bound a ring minimal surface not of minimum type.

Our methods are based on the general critical point theory [11], [8]. We seek a function $W(P)$ defined on a metric space Π and of such a character that its critical points define ring minimal surfaces bounded by g_0 and g_1, or in the "restricted" case (see §7) define disc minimal surfaces bounded respectively by g_0 and g_1. To apply the general theory the function W should be boundedly compact, regular at infinity, and weakly upper-reducible in the sense of [11].

Received January 28, 1941; presented to the American Mathematical Society December 1939. See [10] and [13] for abstracts. Max Shiffman has written an independent paper on problems similar to the ones treated here. The methods he uses differ from those used here.

The restricted case is represented by points on the "restricted" projection Z of Π. In a special sense Z consists of ideal points added to the space of unrestricted points of Π. The metrizing of Π becomes difficult neighboring Z. The function W is obtained by an essential modification of the function defined by the classical Dirichlet integral sum. (See [2], [3], [4], [5].)

We determine a suitable space Π and function $W(P)$, and with the aid of the general theory draw the relevant conclusions.

2. **Monotone transformations.** The space Π on which $W(P)$ is to be defined depends for its definition and metric properties upon the characteristics of the special monotone transformations which we shall employ. Cf. [2], [3], [4], [5]. §2 concerns these transformations in their simplest form. In §3 we study the subclass of restricted transformations. In §4, pairs of such transformations obtainable one from the other by conformal transformations are studied. In §5, the classical Dirichlet integral sums are reviewed. In §6, the space Π and function $W(P)$ are defined.

The material and results of §2 are largely expository in character. We are concerned with transformations of the form

$$\beta = h(\alpha) \qquad (-\infty < \alpha < \infty), \tag{2.1}$$

in which $h(\alpha)$ is real and non-decreasing and

$$h(\alpha + 2\pi) \equiv h(\alpha) + 2\pi. \tag{2.2}$$

The variable α may be regarded as the arc length on a unit circle of points $(u, v) = (\cos \alpha, \sin \alpha)$. The function $h(\alpha)$ then defines a transformation of the unit circle into itself.

The points of discontinuity of $h(\alpha)$ are at most enumerable, and form a discrete set. At such points $h(\alpha)$ has right and left limits. Among transformations $h(\alpha)$ which are discontinuous we shall be principally concerned with those whose discontinuities occur exclusively at a set of points of the form $\alpha \equiv c$ (mod 2π), where c is a constant, with $h(\alpha)$ constant between these points of discontinuity. Such transformations are termed *degenerate*. If a transformation is degenerate, the right and left limits of $h(\alpha)$ at each singularity differ by 2π.

Let $h(\alpha)$ and $k(\alpha)$ be two admissible transformations, and let r be any integer. We introduce the Lebesgue integral

$$I(h, k, r) = \left[\int_0^{2\pi} [h(\alpha) - k(\alpha) + 2\pi r]^2 \, d\alpha\right]^{\frac{1}{2}}. \tag{2.3}$$

The minimum of I, as r ranges over all integers, will be denoted by $[h, k]$. That this function satisfies the triangle axiom

$$[h, k] \leqq [h, l] + [l, k] \tag{2.4}$$

may be seen as follows.

Let r and s be integers such that

$$I(h, l, r) = [h, l], \qquad I(l, k, s) = [l, k],$$

and let $n = r + s$. We see that

$$I(h, k, n) \leqq I(h, l, r) + I(l, k, s). \tag{2.5}$$

The relation (2.4) follows upon noting that $[h, k]$ is at most the left member of (2.5).

Two transformations $h(\alpha)$ and $k(\alpha)$ will be said to be *almost congruent* if there exists an integer r such that

$$h(\alpha) \equiv k(\alpha) + 2\pi r, \tag{2.6}$$

except at most at the singularities of h and k. If h and k are almost congruent, $[h, k] = 0$. Conversely if $[h, k] = 0$, h and k are almost congruent. Two transformations which are almost congruent will be regarded as *representatives of the same point* in a metric space I with distance function $[h, k]$.

If h and k are almost congruent and h or k is continuous, then both h and k are continuous and h and k are congruent mod 2π in the ordinary sense.

Convergence and limits in the space I will be referred to as *I-convergence* and *I-limits* respectively. If a sequence $h^n(\alpha)$ $(n = 1, 2, \cdots)$ of transformations I-converges to $h(\alpha)$, then suitably chosen transformations $k^n(\alpha)$ respectively congruent to the transformations $h^n(\alpha)$ converge *in the mean*[1] on the interval $(0, 2\pi)$ to $h(\alpha)$. That is,

$$\lim_{n=\infty} \int_0^{2\pi} [k^n(\alpha) - h(\alpha)]^2 \, d\alpha = 0.$$

A sequence $k^n(\alpha)$ $(n = 1, 2, \cdots)$ of transformations such that

$$\lim_{n=\infty} k^n(\alpha) = k(\alpha)$$

at each point α on a set ω will be said to converge *pointwise* to $k(\alpha)$ on ω.

If the sequence of transformations $k^n(\alpha)$ converges almost everywhere to $k(\alpha)$, it follows from the Lebesgue integration theory that $k^n(\alpha)$ converges in the mean to $k(\alpha)$. Conversely if $k^n(\alpha)$ converges in the mean to $k(\alpha)$, $k^n(\alpha)$ will converge pointwise to $k(\alpha)$ at each point of continuity of $k(\alpha)$. This follows from the monotone character of our transformations. We thus have the following lemma.

LEMMA 2.1. *A necessary and sufficient condition that a sequence of transformations converge in the mean to a transformation $k(\alpha)$ is that the sequence converge pointwise to $k(\alpha)$ except at most at the singular points of $k(\alpha)$.*

We continue with the following lemma.

LEMMA 2.2. *The space I is compact.*

[1] Convergence in the mean shall always refer to the interval $(0, 2\pi)$.

Let $h^n(\alpha)$ $(n = 1, 2, \cdots)$ be a sequence of transformations. We seek a limit transformation of this sequence on the space I. It follows from the definition of distances $[h, k]$ on I that the transformations $h^n(\alpha)$ may be replaced by congruent transformations without loss of generality. In particular, we can suppose that the functions $h^n(\alpha)$ are bounded on the interval $(0, 2\pi)$. According to a theorem of Helly (cf. [15], p. 17) there then exists a subsequence $k^m(\alpha)$ $(m = 1, 2, \cdots)$ of the sequence $h^n(\alpha)$ $(n = 1, 2, \cdots)$ which converges pointwise to a transformation $k(\alpha)$. Hence the sequence $k^m(\alpha)$ converges in the mean to $k(\alpha)$.

The lemma follows.

Convergence in the mean of a sequence of transformations to a transformation $k(\alpha)$ implies pointwise convergence except at the points of discontinuity of $k(\alpha)$, but it does not imply uniform convergence unless $k(\alpha)$ is continuous. We have the following lemma.

Lemma 2.3. *If the limit in the mean of a sequence of transformations is continuous, the convergence is uniform with respect to* α.

This lemma follows at once from the monotone character of our transformations.

3. **Restricted transformations.** We shall choose three values of α, $\alpha_1 < \alpha_2 < \alpha_3$, on the interval $0 \leqq \alpha < 2\pi$, and consider *continuous* transformations of the type admitted in §2, subject to the conditions

$$\varphi(\alpha_i) = \alpha_i \qquad (i = 1, 2, 3). \tag{3.1}$$

Such transformations will be termed *restricted*. Restricted transformations will be designated by Greek letters φ, ψ, etc., while unrestricted transformations will be designated by Roman letters h, k, etc. as before. We term (3.1) the *fixed point condition*. The values α_i will be fixed throughout the paper.

We shall be concerned with directly conformal transformations of the unit disc $u^2 + v^2 < 1$ onto itself. Such transformations may be continuously extended to the circle $u^2 + v^2 = 1$, and when so extended, these transformations are one-to-one. On the unit circle a point $(\cos \alpha, \sin \alpha)$ is replaced by a point $(\cos \beta, \sin \beta)$ where β is obtainable, mod 2π, from α by a transformation

$$\beta = T(\alpha)$$

of the type admitted in §2. We term $T(\alpha)$ a *Möbius transformation* of a unit circle, or simply a Möbius transformation. The function $T(\alpha)$ is strictly increasing. The given conformal transformation determines not only $T(\alpha)$ but each congruent image of $T(\alpha)$. We shall write product transformations of the form $h[T(\alpha)]$ more simply as hT.

If a transformation $h(\alpha)$ has the form φT, where φ is a restricted transformation and T a Möbius transformation, φT will be termed a *canonical representation* of h.

LEMMA 3.1. *A continuous transformation* $h(\alpha)$ *of* §2 *admits at least one canonical representation* φT.

The values α_i used in defining the fixed point condition will be assumed by $h(\alpha)$ at least once, since $h(\alpha)$ is continuous and varies through all real values. Thus there exist values c_i $(i = 1, 2, 3)$ such that

$$\alpha_i = h(c_i) \qquad (i = 1, 2, 3), \tag{3.2}$$

with

$$c_1 < c_2 < c_3 < c_1 + 2\pi.$$

There will accordingly exist a Möbius transformation $S(\alpha)$ (see [1], p. 6) such that

$$c_i = S(\alpha_i) \qquad (i = 1, 2, 3). \tag{3.3}$$

With this choice of S,

$$h[S(\alpha_i)] = h(c_i) = \alpha_i, \tag{3.4}$$

so that hS is a restricted transformation φ. Hence $h = \varphi S^{-1}$, and the proof of the lemma is complete.

If the transformation $h(\alpha)$ of Lemma 3.1 is strictly increasing, $h(\alpha)$ admits but one canonical representation $h = \varphi T$. For in this case h has a single-valued inverse, and

$$T^{-1}(\alpha_i) = h^{-1}\{\varphi(\alpha_i)\} = h^{-1}(\alpha_i) \qquad (i = 1, 2, 3).$$

Thus T is thereby uniquely determined.

We shall need the following lemma.

LEMMA 3.2. *Let* φ_n *and* h_n $(n = 1, 2, \cdots)$ *be sequences of transformations converging pointwise to transformations* ψ *and* k *respectively. If* ψ *is continuous,* $\varphi_n h_n$ *converges pointwise to* ψk.

Were ψ not continuous one could show by example that $\varphi_n h_n$ would not necessarily converge as stated to ψk.

Since ψ is continuous, φ_n converges uniformly to ψ, and we see that $\varphi_n h_n$ converges pointwise to ψk. The lemma also follows readily for convergence on I.

LEMMA 3.3. *Lemma* 3.2 *holds if pointwise convergence is replaced throughout by convergence on* I.

4. **Möbius transformations.** Let z and w be complex variables. The most general directly conformal transformation of the unit circular disc into itself has the form (see [1], p. 18)

$$w = e^{ib} \frac{a - z}{1 - \bar{a}z}, \tag{4.1}$$

where a is a complex constant, $\bar{a}$ its conjugate, $|a| < 1$, and b is a real constant. The corresponding Möbius transformations, $\beta = T(\alpha)$, of the unit circle $|z| = 1$ into the unit circle $|w| = 1$ are obtained (mod 2π) upon setting

$$z = e^{i\alpha}, \qquad w = e^{i\beta} \tag{4.2}$$

in (4.1). The functions $T(\alpha)$ may be regarded as depending on the parameters a and b, and they admit representations by functions $T(\alpha, a, b)$ which vary continuously with (α, a, b), provided $|a| < 1$. In particular, if (a, b) tends to a pair (a^*, b^*) for which $|a^*| < 1$, a branch function $T(\alpha, a, b)$ will converge uniformly to $T(\alpha, a^*, b^*)$.

If $|a| = 1$, the transformation (4.1) is singular. Here $a\bar{a} = 1$, and the denominator in (4.1) vanishes when $z = a$. When $|a| = 1$ and $z \neq a$, (4.1) takes the form

$$w = ae^{ib}. \tag{4.3}$$

Upon making the substitution (4.2) and setting $a = e^{ic}$, we find that for $\alpha \not\equiv c$ (mod 2π),

$$\beta \equiv b + c \qquad (\text{mod } 2\pi). \tag{4.4}$$

Singular transformations such as this are not included among our Möbius transformations.

LEMMA 4.1. *An I-limit $h(\alpha)$ of a sequence of Möbius transformations is a Möbius transformation or degenerate.*

Let $T_n(\alpha)$ $(n = 1, 2, \cdots)$ be a sequence of Möbius transformations with an I-limit $h(\alpha)$, and let (a_n, b_n) be the constants (a, b) of the transformation (4.1) defining $T_n(\alpha)$. The pairs (a_n, b_n) will have at least one limit pair (a^*, b^*). Without loss of generality we can suppose that the sequence (a_n, b_n) converges to (a^*, b^*). There are two cases to consider.

Case I. $|a^*| < 1$. Among the various congruent choices of $T(\alpha, a, b)$, choose $T(\alpha, a, b)$ so that it is continuous in its arguments for (a, b) on a neighborhood of (a^*, b^*). Now $T_n(\alpha)$ is congruent to $T(\alpha, a_n, b_n)$ and the latter converges uniformly with respect to α to $T(\alpha, a^*, b^*)$ as n becomes infinite. Hence $T_n(\alpha)$ I-converges to $T(\alpha, a^*, b^*)$. In Case I then $T_n(\alpha)$ I-converges to a Möbius transformation.

Case II. $|a^*| = 1$. When $a = a^*$, the transformation (4.1) implies the relation (4.4). We see that as n becomes infinite $T(\alpha, a_n, b_n)$ converges pointwise to a transformation T satisfying (4.4) except at most when $\alpha \equiv c$ (mod 2π). The transformation T is clearly degenerate, and the I-limit of the sequence $T_n(\alpha)$.

The proof of the lemma is complete.

The space I is compact. The preceding lemma accordingly leads to the following lemma.

LEMMA 4.2. *The set of Möbius transformations T and degenerate transformations forms a compact subset of I.*

5. **The functions $D(p)$ and $D(p_0, p_1, \rho)$.** Let x be a vector representing a point $(x_1, \cdots, x_n)$ in a Euclidean space E of n rectangular coördinates x_i. In E let p be a simple closed curve represented by the vector[2] equation

$$x = p(\alpha) \qquad (0 \leqq \alpha \leqq 2\pi). \tag{5.1}$$

We suppose $p(\alpha)$ such that $p(\alpha + 2\pi) \equiv p(\alpha)$. Let (r, θ) be polar coördinates in a plane of rectangular coördinates (u, v). The representation $p(\alpha)$ *defines* a "harmonic surface" S of the form $x = x(u, v)$ in which $x(u, v)$ is harmonic for $r < 1$, continuous for $r \leqq 1$, and such that

$$x(\cos \theta, \sin \theta) \equiv p(\theta).$$

Let $\omega(a)$ represent the region $u^2 + v^2 < a^2 < 1$. Set

$$D(p) = \tfrac{1}{2} \lim_{a=1} \iint_{\omega(a)} \left[\left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial x}{\partial v}\right)^2\right] du\, dv, \tag{5.2}$$

understanding that $D(p)$ may be infinite.

We turn to the case of two contours. Let

$$x = p_0(\alpha), \qquad x = p_1(\alpha) \qquad (0 \leqq \alpha \leqq 2\pi) \tag{5.3}$$

be a representation of two simple non-intersecting closed curves, with $p_0(\alpha + 2\pi) \equiv p_0(\alpha)$ and $p_1(\alpha + 2\pi) \equiv p_1(\alpha)$. Let B be a region in the (u, v)-plane bounded by concentric circles C_0 and C_1, with centers at the origin and radii σ_0 and σ_1 respectively. We suppose that $\sigma_0 < \sigma_1$. Let $x = H(u, v)$ be a vector representing a ring harmonic surface defined as follows. The function $H(u, v)$ shall be harmonic on B, continuous on $\bar{B}$, the closure of B, and such that

$$H(\sigma_k \cos \theta, \sigma_k \sin \theta) \equiv p_k(\theta) \qquad (k = 0, 1).$$

Let r_0 and r_1 be constants such that

$$\sigma_0 < r_0 < r_1 < \sigma_1 .$$

Let $\rho = \sigma_0/\sigma_1$ and set

$$D(p_0, p_1, \rho) = \lim_{r_k = \sigma_k} \frac{1}{2} \iint_{r_0 < r < r_1} \left[\left(\frac{\partial H}{\partial u}\right)^2 + \left(\frac{\partial H}{\partial v}\right)^2\right] du\, dv \quad (k = 0, 1), \tag{5.4}$$

understanding that $D(p_0, p_1, \rho)$ may be infinite. The limiting integral here defined depends merely on the ratio ρ and not on σ_0 and σ_1 in any other way, since a simple radial magnification of the region B obviously will lead to the same limiting integral (5.4).

Let

$$a_k(m), \qquad b_k(m) \qquad (k = 0, 1)$$

[2] Points, curves and surfaces will be represented in vector form in the space E unless otherwise stated.

be the m-th vector Fourier coefficients of $p_k(\theta)$. When $D(p_0, p_1, \rho)$ is finite, it is known[3] ([7], p. 280) that

$$D(p_0, p_1, \rho) = D(p_0) + D(p_1) + T(p_0, p_1, \rho) + R(p_0, p_1, \rho), \tag{5.5}$$

where, on summing with respect to m,

$$\begin{aligned} R(p_0, p_1, \rho) = \pi \sum \frac{m\rho^{2m}}{1-\rho^{2m}} [a_0^2(m) + b_0^2(m) + a_1^2(m) + b_1^2(m)] \\ -2\pi \sum \frac{m\rho^m}{1-\rho^{2m}} [a_0(m)a_1(m) + b_0(m)b_1(m)] \quad (m = 1, 2, \cdots), \end{aligned} \tag{5.6}$$

$$T(p_0, p_1, \rho) = -\frac{\pi}{4} \frac{[a_1(0) - a_0(0)]^2}{\log \rho} \qquad (0 < \rho < 1). \tag{5.7}$$

We shall extend the definitions (5.6) and (5.7) by setting

$$T(p_0, p_1, 0) = R(p_0, p_1, 0) = 0. \tag{5.8}$$

We shall remark on *the convergence and continuity* of the functions T and R. Let M be a constant such that

$$| p_k(\theta) | < \tfrac{1}{2}M \qquad (k = 0, 1). \tag{5.9}$$

The Fourier coefficients $a_k(m)$, $b_k(m)$ then have magnitudes at most M. Moreover, these coefficients vary continuously with $p_0(\theta)$ and $p_1(\theta)$. In defining this continuity we think of the sets (p_0, p_1, ρ) as points in a metric space in which the distance to a second point (q_0, q_1, σ) is the square root of

$$\frac{1}{2\pi}\int_0^{2\pi} \{| p_0(\theta) - q_0(\theta) |^2 + | p_1(\theta) - q_1(\theta) |^2\} \, d\theta + (\sigma - \rho)^2.$$

It appears at once that $T(p_0, p_1, \rho)$ is continuous.

To see that $R(p_0, p_1, \rho)$ is continuous, let ρ be restricted to an interval

$$0 \leqq \rho < \mu < 1, \tag{5.10}$$

where μ is a constant. Under the conditions (5.9) and (5.10) the sum of the two terms in (5.6) involving m has a magnitude less than

$$\frac{cm\mu^m}{1-\mu}, \tag{5.11}$$

where c is a constant independent of m. The series of terms (5.11) converges. Hence the series (5.6) converges uniformly with respect to ρ and with respect to any arguments on which the Fourier coefficients continuously depend, (5.9) holding. We draw the following conclusions.

LEMMA 5.1. *The functions $T(p_0, p_1, \rho)$ and $R(p_0, p_1, \rho)$ are continuous in their arguments and are bounded subject to* (5.9) *and* (5.10).

[3] This was shown by Douglas in [4], p. 339.

We add the following lemma.

LEMMA 5.2. *For a fixed ρ $(0 < \rho < 1)$, $D(p_0, p_1, \rho)$ is finite if and only if $D(p_0)$ and $D(p_1)$ are finite.*

To verify this lemma, replace p_k $(k = 0, 1)$ by the curve

$$H(r_k \cos \theta, r_k \sin \theta) = q_k(\theta) \qquad (k = 0, 1),$$

where $\sigma_0 < r_0 < r_1 < \sigma_1$. Set $\tau = r_0/r_1$. Then $D(q_0, q_1, \tau)$ is finite and (5.5) holds in the form

$$(5.12) \quad D(q_0, q_1, \tau) = D(q_0) + D(q_1) + T(q_0, q_1, \tau) + R(q_0, q_1, \tau).$$

Let r_k tend to σ_k $(k = 0, 1)$. The terms T and R in (5.12) remain bounded, while the functions

$$D(q_0, q_1, \tau), \qquad D(q_0), \qquad D(q_1)$$

tend to the functions

$$D(p_0, p_1, \rho), \qquad D(p_0), \qquad D(p_1),$$

respectively by definition. The lemma follows.

We shall have occasion to differentiate $T(p_0, p_1, \rho)$ and $R(p_0, p_1, \rho)$ with respect to certain parameters.

For $\rho \neq 0$, T_ρ exists and becomes positively infinite as ρ tends to 0, provided

$$| a_1(0) - a_0(0) | \neq 0.$$

Moreover, R_ρ is readily seen to exist and to be obtainable by termwise differentiation of the series in (5.6). Subject to (5.9) and (5.10) R_ρ is bounded, and subject to (5.9) it tends to 0 uniformly with respect to p_0 and p_1 as ρ tends to 0.

Let $\beta = A(\alpha, e)$ be an analytic family of analytic transformations of α of the type of §2 in which e ranges on an interval $0 \leqq e \leqq e_1$, on which $A_\alpha(\alpha, e) \neq 0$. We shall set

$$R(p_0, p_1, \rho) + T(p_0, p_1, \rho) = \Omega(p_0, p_1, \rho),$$

and

$$\Omega[p_0(A), p_1(A), \rho] = f(p_0, p_1, e, \rho).$$

We shall prove the following lemma.

LEMMA 5.3. *The partial derivatives f_e exist and are continuous in their arguments (p_0, p_1, e, ρ) for arbitrary admissible representations p_0, p_1, for $0 \leqq e \leqq e_1$, and for $0 \leqq \rho < 1$.*

It will be sufficient to prove the lemma for uniformly bounded $| p_0(\alpha) |$ and $| p_1(\alpha) |$ and for $0 \leqq \rho < \mu < 1$, where μ is a positive constant. The lemma will follow from Lemma 4.1 of [7] once we have proved that the Fourier vector

coefficients $a_k(m)$, $b_k(m)$ of $p_k(A)$ $(k = 0, 1)$ satisfy conditions of the form

$$\left| \frac{\partial a_k(m)}{\partial e} \right| < mK, \qquad \left| \frac{\partial b_k(m)}{\partial e} \right| < mK, \tag{5.13}$$

where K is a positive constant independent of p_k, ρ, and e.

We have

$$b_k(m) = \frac{1}{2\pi} \int_0^{2\pi} p_k[A(\alpha, e)] \sin m\alpha \, d\alpha. \tag{5.14}$$

Set $t = A(\alpha, e)$. The inverse transformation $\alpha = \alpha(t, e)$ is of the type of §2 and is analytic for all real t and for $0 \leqq e \leqq e_1$. From (5.14) it follows that

$$b_k(m) = \frac{1}{2\pi} \int_0^{2\pi} p_k(t) \sin [m\alpha(t, e)]\alpha_t(t, e) \, dt. \tag{5.15}$$

We can differentiate under the integral sign with respect to e. One can treat $a_k(m)$ similarly. Relations (5.13) follow for bounded $| p_k(\alpha) |$, for $0 \leqq \rho < \mu < 1$, and for $0 \leqq e \leqq e_1$. The lemma follows from Lemma 4.1 of [7].

We turn to the case of one contour. Let $\lambda(\alpha)$ be a real analytic function with a period 2π, and let $p(\alpha)$ be an admissible representation of a simple closed curve g. Let e_1 be a positive constant so small that $1 + e\lambda'(\alpha) \neq 0$ for $0 \leqq e \leqq e_1$. The following lemma is a consequence of Lemma 6.2 of [11].

Lemma 5.4. *For $0 \leqq e \leqq e_1$ and for $D(p)$ finite, the e-derivative $S(p, e)$ of $D\{p[\alpha + e\lambda(\alpha)]\}$ exists. If c is a finite constant, and q a representation for which $S(q, 0) < 0$, then $S(p, e)$ is negative and bounded from 0 for e sufficiently near 0 and for p on a sufficiently small neighborhood of q relative to the set on which $D(p) \leqq c$.*

In [11] neighborhoods in the space of representations p are not defined by "convergence in the mean". However, if they are so defined, the above lemma remains valid. In fact, with this change in the notion of convergence $S(p, 0)$ remains continuous in p, as one readily sees from (6.17) in [11], using the fact that the integral (6.17) has a removable singularity. The remaining analysis goes as before.

A deformation. If $h_r(\alpha)$ and $k_r(\alpha)$ $(r = 0, 1)$ are arbitrary transformations of I, we shall have occasion to consider a deformation of $h_r(\alpha)$ into $k_r(\alpha)$ in which $h_r(\alpha)$ is replaced at the time t by

$$h_r^t = th_r + (1 - t)k_r \qquad (r = 0, 1; 0 \leqq t \leqq 1). \tag{5.16}$$

For fixed representations $p_0(\alpha)$ and $p_1(\alpha)$, we set

$$\Omega[p_0(h_0^t), p_1(h_1^t), \rho] = F(h_0, h_1, t, \rho).$$

It is easily seen that F is continuous in its arguments. We continue with the following lemma.

LEMMA 5.5. *If the representations* $p_r(\alpha)$ $(r = 0, 1)$ *are of class* C^1, *the t-derivative of the function* $F(h_0, h_1, t, \rho)$ *exists and is continuous for* (h_0, h_1) *on* I^2, *for* $0 \leq t \leq 1$, *and for* $0 \leq \rho < 1$. *Moreover, for* ρ *bounded from* 1, F_t *tends to* 0 *uniformly as* (h_0, h_1) *tends to* (k_0, k_1) *on* I^2.

The Fourier coefficient $b_r(m)$ of $p_r(h_r^t)$ has the form

$$b_r(m) = \frac{1}{2\pi}\int_0^{2\pi} p_r(h_r^t) \sin m\alpha \, d\alpha \qquad (r = 0, 1),$$

so that

$$\frac{\partial}{\partial t} b_r(m) = \frac{1}{2\pi}\int_0^{2\pi} (h_r - k_r)p_r'(h_r^t) \sin m\alpha \, d\alpha.$$

A similar formula holds for the t-derivative of $a_r(m)$. The lemma follows from the representation of $\Omega(p_0, p_1, \rho)$ by means of (5.6) and (5.7).

6. **The functions** $A(g, h)$ **and** $B(h_0, h_1, \rho)$. Let g be a simple rectifiable curve with a vector representation of the form

$$x = g(t) \qquad (0 \leq t < 2\pi),$$

with t proportional to the arc length and $g(t + 2\pi) \equiv g(t)$. We shall use other representations of g of the form

$$x = p(\alpha) = g[h(\alpha)],$$

where $h(\alpha)$ is a continuous transformation of the type admitted in §2. According to Douglas, the Dirichlet sum $D(p)$ equals the improper integral

$$A(g, h) = \frac{1}{16\pi}\iint_Q \frac{[p(\alpha) - p(\beta)]^2}{\sin^2[\frac{1}{2}(\alpha - \beta)]} d\alpha \, d\beta \tag{6.1}$$

extended over the parallelogram

$$Q\colon\ 0 \leq \alpha \leq 2\pi, \quad -\pi < \alpha - \beta \leq \pi, \tag{6.2}$$

excluding the singular line $\alpha = \beta$. In case $h(\alpha)$ is discontinuous we shall understand that $A(g, h)$ is defined by (6.1).

In the case of two contours we shall start with two curves g_0 and g_1 satisfying the following conditions:

(6a) *The curve* g_k *shall be a simple, rectifiable, closed curve with the representation*

$$x = g_k(t) \qquad (k = 0, 1; 0 \leq t < 2\pi),$$

where t *is proportional to the arc length and* $g_k(t + 2\pi) = g_k(t)$.

(6b) *The derived vectors* g_k' *shall satisfy Lipschitz conditions*

$$| g_k'(t + \Delta t) - g_k'(t) | < M \, | \Delta t |,$$

where M *is a positive constant.*

The condition (6b) will subsequently be replaced by the chord arc condition.

The function $B(h_0, h_1, \rho)$. The curve $p_k(\theta)$ of the preceding section will now be replaced by the curve

$$p_k(\theta) = g_k[h_k(\theta)] \qquad (k = 0, 1), \tag{6.3}$$

where $h_k(\theta)$ is a continuous transformation of the type admitted in §2. Subject to (6.3) we set

$$D(p_0, p_1, \rho) = B(h_0, h_1, \rho) \qquad (0 < \rho < 1), \tag{6.4}$$

and write (5.5) in the form

$$B(h_0, h_1, \rho) = A(g_0, h_0) + A(g_1, h_1) + \Omega(p_0, p_1, \rho). \tag{6.5}$$

When either h_0 or h_1 is discontinuous, or $\rho = 0$, we understand that $B(h_0, h_1, \rho)$ is defined by (6.5) with $\Omega = T + R$ given by (5.6) and (5.7), subject to (6.3). The function $B(h_0, h_1, \rho)$ is defined for $0 \leqq \rho < 1$ and for h_0 and h_1 arbitrary transformations of the type admitted in §2. Nevertheless it is not adequate for our purposes.

To apply the general critical point theory it is desirable that the function F whose critical points are sought be boundedly compact, regular at infinity, and weakly upper-reducible. See §2, [11]. Moreover, it is desirable that each homotopic critical point define a minimal surface. This is not the case with $B(h_0, h_1, \rho)$.

Some of the difficulties inherent in the problem have already appeared in the study of $A(g, h)$ in the case of one contour g. The transformations h were restricted transformations. Without this limitation the following difficulties would be encountered in the study of $A(g, h)$.

(1) The subspace of continuous transformations h of I, for which $A(g, h)$ is at most a finite constant, is *not* in general *compact* unless completed by the degenerate transformations.

(2) If $h(\alpha)$ defines a minimal surface, the transformation hT, where T is an arbitrary Möbius transformation, defines the *same* minimal surface (apart from parametrization).

(3) If h is degenerate, $A(g, h) = 0$. Such transformations h give a *trivial minimum* to $A(g, h)$.

(4) The space I contains *non-bounding* 1-cycles composed of 1-parameter families of transformations $h(\alpha)$ of the form

$$h(\alpha) = k(\alpha) + \tau \qquad (0 \leqq \tau < 2\pi),$$

where $k(\alpha)$ is a particular transformation, and τ is a parameter which varies from transformation to transformation.

The transformations $k(\alpha)$ and $k(\alpha) + 2\pi$ are to be regarded as defining the same point on I in accordance with our conventions, thereby closing the 1-cycle.

(5) The function $A(g, h)$ is *not upper-reducible* in h at degenerate points h.

The domain of definition of $B(h_0, h_1, \rho)$ may be taken as the product space

$I \times I \times J$, where J is the interval $0 \leqq \rho < 1$. The subspace on which $\rho = 0$ is the product $I \times I$, and on this subspace

$$B(h_0, h_1, 0) = A(g_0, h_0) + A(g_1, h_1).$$

The difficulties (1) to (5) just enumerated appear in corresponding form in this subspace.

We shall surmount the difficulties by the introduction of a new space Π of points P, and a new function $W(P)$ obtained from $B(h_0, h_1, \rho)$ by essential modifications.

7. **The space Π and function $W(P)$.** We shall make use of the possibility of representing continuous transformations $h(\alpha)$ in the canonical form φT by taking restricted transformations and Möbius transformations as new independent variables. The procedure is as follows.

Let J denote the interval $0 \leqq \rho < 1$. Let $I^4 \times J$ denote the product of I^4 by J. The space Π shall consist of points

$$P = (\varphi_0, \varphi_1, \tau_0, \tau_1, \rho),$$

in which φ_k $(k = 0, 1)$ is a restricted transformation, τ_k a Möbius or degenerate transformation, and ρ is on J. We add the following conventions of identity. The pair $\{\varphi_0, \varphi_1\}$ enclosed in braces will be termed the *restricted projection* of P, and will be regarded as a point on I^2. For $\rho = 0$, two points with the same restricted projection shall be identical on Π and may be represented by their restricted projection $\{\varphi_0, \varphi_1\}$. For $\rho > 0$, the conventions of identity on Π are those on $I^4 \times J$. No point of Π for which $\rho > 0$ shall be identified with a point for which $\rho = 0$.

The distance function PQ will now be defined. Let

$$Q = (\psi_0, \psi_1, \sigma_0, \sigma_1, a)$$

be a second point of Π. Let $d(P, Q)$ be the distance between P and Q given by the conventional metric[4] of $I^4 \times J$. Let

$$\delta(P, Q) = [\varphi_0, \psi_0] + [\varphi_1, \psi_1] + \rho + a.$$

Finally let

$$PQ = \min [d(P, Q), \delta(P, Q)]. \tag{7.1}$$

We continue with a proof of the following lemma.

[4] Let S be the product of metric spaces S_i $(i = 1, \cdots, m)$ with points p_i, q_i, etc., and distance functions $\delta_i(p_i, q_i)$. The distance between points

$$p = (p_1, \cdots, p_m), \qquad q = (q_1, \cdots, q_m),$$

of S will be conventionally defined as the sum

$$pq = \delta_1(p_1, q_1) + \cdots + \delta_m(p_m, q_m).$$

LEMMA 7.1. *The distance $PQ = 0$ if and only if $P = Q$ on* Π.

We begin by assuming that $PQ = 0$. Two cases arise.

Case I. $\delta(P, Q) = 0$. This is possible only if $\rho = a = 0$, and if the "restricted projections" of P and Q on I^2 are identical. Hence $P = Q$.

Case II. $\delta(P, Q) \neq 0$. In this case $d(P, Q) = 0$. In particular $\rho = a$. Moreover, ρ and a are positive. Otherwise $d(P, Q) \geqq \delta(P, Q)$, so that $\delta(P, Q) = 0$, contrary to hypothesis. With $\rho = a > 0$, the condition $d(P, Q) = 0$ implies that P and Q are identical in the sense of points on $I^4 \times J$. Thus $P = Q$ on Π.

Conversely, it is seen that when $P = Q$, $PQ = 0$, and the proof of the lemma is complete.

LEMMA 7.2. *The distance function satisfies the triangle relation*

$$PP'' \leqq PP' + P'P''.$$

Set

$$\begin{aligned} P \;\; &= (\varphi_0, \varphi_1, h_0, h_1, \rho), \\ P' \;&= (\varphi_0', \varphi_1', h_0', h_1', \rho'), \\ P'' &= (\varphi_0'', \varphi_1'', h_0'', h_1'', \rho''). \end{aligned}$$

By virtue of (7.1)

$$PP'' \leqq d(P, P'') \leqq d(P, P') + d(P', P''), \tag{7.2}$$

$$PP'' \leqq \delta(P, P'') \leqq \delta(P, P') + \delta(P', P''). \tag{7.3}$$

Let the projection $\{\varphi_0, \varphi_1\}$ of P on I^2 be denoted by p, those of P' and P'' by p' and p'' respectively. Let distances on I^2 be denoted by $|p, p'|$, etc. Then

$$\begin{aligned} PP'' &\leqq \delta(PP'') = |p, p''| + \rho + \rho'' \\ &\leqq |p, p'| + |p', p''| + |\rho - \rho'| + \rho' + \rho'' \\ &\leqq \{|p, p'| + |\rho - \rho'|\} + \{|p', p''| + \rho' + \rho''\} \leqq d(P, P') + \delta(P', P''). \end{aligned}$$

A similar relation results upon interchanging P and P''. Thus we have the two relations

$$PP'' \leqq d(P, P') + \delta(P', P''), \tag{7.4}$$

$$PP'' \leqq \delta(P, P') + d(P', P''). \tag{7.5}$$

The relation to be proved has the form

$$PP'' \leqq \min\,[d(P, P'), \delta(P, P')] + \min\,[d(P', P''), \delta(P', P'')].$$

This relation reduces to one of the relations (7.2) to (7.5), depending on what the minima indicated in the formula in fact are.

The distance function PQ thus satisfies the usual metric relations.

The subspace $\rho = 0$ of Π may be identified with the subspace of restricted pairs $\{\varphi_0, \varphi_1\}$ of the space I^2 both as to points and as to metric. Understanding that a point $(\varphi_0, \varphi_1, h_0, h_1, \rho)$ of Π shall correspond to the point of $I^4 \times J$ with the same components, we shall show that for the subspace $\rho > 0$ of Π and the corresponding subspace of $I^4 \times J$ this correspondence is topological. For if P_0 is a fixed point of Π for which $\rho > 0$, and if P_0P is sufficiently small,

$$P_0P < \delta(P_0, P),$$

so that

$$P_0P = d(P_0, P).$$

Thus sufficiently small spherical neighborhoods of P_0 on Π are identical with the corresponding neighborhoods on $I^4 \times J$.

A point P for which h_0 or h_1 is degenerate and $\rho > 0$ will be termed *degenerate*.

The function $W(P)$. For each point

$$P = (\varphi_0, \varphi_1, \tau_0, \tau_1, \rho) \tag{7.6}$$

of Π, we set

$$W(P) = A(g_0, \varphi_0) + A(g_1, \varphi_1) + \Omega[g_0(\varphi_0\tau_0), g_1(\varphi_1\tau_1), \rho], \tag{7.7}$$

where $\Omega = R + T$ and $g_k(\varphi_k\tau_k)$ is an abbreviation for

$$g_k\{\varphi_k[\tau_k(\alpha)]\} = p_k(\alpha) \qquad (k = 0, 1). \tag{7.8}$$

We shall compare $W(P)$ with $D(p_0, p_1, \rho)$ where p_k is given by (7.8), and D by $B(h_0, h_1, \rho)$ in (6.5).

When P is non-degenerate and $\rho > 0$,

$$A(g_k, \varphi_k\tau_k) = A(g_k, \varphi_k), \tag{7.9}$$

since τ_k is a Möbius transformation. In this case

$$W(P) = D(p_0, p_1, \rho).$$

If P is degenerate, at least one of the left members of (7.9) is null, so that

$$W(P) > D(p_0, p_1, \rho).$$

Finally, for $\rho = 0$,

$$W(P) = A(g_0, \varphi_0) + A(g_1, \varphi_1) \geqq D(p_0, p_1, \rho).$$

Thus, in all cases,

$$W(P) \geqq D(p_0, p_1, \rho), \tag{7.10}$$

for p_k given by (7.8).

We continue with the following lemma.

LEMMA 7.3. *When ρ tends to 1, $W(P)$ becomes infinite uniformly with respect to the remaining components of P.*

Let ρ tend to 1, holding the remaining components of P fast. When the transformations $h_k(\alpha)$ are continuous or degenerate, $D[g_0(h_0), g_1(h_1), \rho]$ becomes infinite uniformly as ρ tends to 1, in accordance with a formula of Lebesgue explicitly developed for this purpose by Courant. See [3], Lemma C, p. 701. But (7.10) holds in all cases, and we infer the truth of the lemma.

8. **Bounded compactness and regularity at infinity.** *Bounded compactness.* Let the subspace of Π on which $W \leqq c$ be denoted by W_c. We shall show that W_c is compact for each positive constant c.

The function $A(g, h)$ is lower semi-continuous as is well known. The function $W(P)$ as defined by (7.7) is likewise lower semi-continuous.

The function $\Omega(p_0, p_1, \rho)$ is continuous in its arguments by virtue of Lemma 5.1. If $[\varphi_k, \tau_k]$ converges on I^2 to a pair $[\psi_k, \sigma_k]$, where ψ_k is restricted, $\varphi_k\tau_k$ will converge on I to $\psi_k\sigma_k$, in accordance with Lemma 3.2. The representation $g_k(\varphi_k\tau_k)$ will then converge in the mean to $g_k(\psi_k\sigma_k)$. The continuity of the last term in (7.7) as a function of P then follows from the continuity of $\Omega(p_0, p_1, \rho)$, at least for $\rho > 0$. When ρ tends to 0, the last term in (7.7) tends to 0 uniformly with respect to the remaining variables in P. Hence $W(P)$ is lower semi-continuous.

For P on W_c, ρ is bounded from 1 in accordance with Lemma 7.3. Hence the term Ω in (7.7) is bounded for P on W_c. It follows from (7.7) that $A(g_k, \varphi_k)$ $(k = 0, 1)$ is likewise bounded for P on W_c. If

$$P^n = (\varphi_0^n, \varphi_1^n, \tau_0^n, \tau_1^n, \rho^n) \qquad (n = 1, 2, \cdots)$$

is a sequence of points on W_c, φ_k^n will admit a subsequence converging to a restricted transformation ψ_k since $A(g_k, \varphi_k^n)$ is bounded. (Cf. [11].) The subspace of Möbius and degenerate transformations τ is compact on I. Hence a subsequence of P_n $(n = 1, 2, \cdots)$ will converge to a point Q of Π. Since $W(P)$ is lower semi-continuous, Q will lie on W_c. Hence Π is boundedly compact relative to $W(P)$.

We say that $W(P)$ is *regular at infinity* on Π if corresponding to each compact subset A of Π on which W is finite, there exists a continuous deformation of A into a set on which W is bounded, a deformation such that any subset of A on which W is bounded is deformed through a set of points on which F is bounded.

Before verifying the condition of regularity at infinity, we recall the definition of the improper integral $H(\theta)$ (see [11], p. 448),

$$H(\theta) = \frac{1}{16\pi} \iint_Q \frac{|\theta(\alpha) - \theta(\beta)|^2}{\sin^2 [\frac{1}{2}(\alpha - \beta)]} d\alpha\, d\beta, \tag{8.1}$$

where $\theta(\alpha)$ is a restricted transformation. It follows from condition (6b) on g_k that g_k satisfies the chord arc condition of the introduction. From this we can infer that $H(\theta)$ and $A(g_k, \theta)$ lie between two positive constant multiples each of the other, where the constants involved are independent of $\theta(\alpha)$.

The deformation $\Delta_\psi(\varphi)$. Corresponding to an arbitrary restricted transformation ψ, we shall define a continuous deformation $\Delta_\psi(\varphi)$ of each restricted transformation φ. In this deformation the time t shall run from 0 to 1 inclusive, and φ shall be replaced at the time t by the transformation

$$\varphi^t(\alpha) = t\psi(\alpha) + (1 - t)\varphi(\alpha). \tag{8.2}$$

The transformation $\varphi^t(\alpha)$ is clearly restricted. The transformation $\varphi(\alpha)$ is thereby deformed into the transformation $\psi(\alpha)$. It follows as in [11], p. 451, that a set of restricted transformations on which $A(g_k, \varphi)$ $(k = 0, 1)$ is bounded is deformed through a set of transformations φ^t on which $A(g_k, \varphi^t)$ is bounded.

The deformation $\Delta_0(P)$. We shall deform Π on itself into a point. We shall let ρ decrease to 0, holding the other components of $P = (\varphi_0, \varphi_1, h_0, h_1, \rho)$ fast. In this deformation the time t shall equal $-\rho$, and shall increase on the interval $-1 < t \leqq 0$. A given point P shall be fixed until t increases to $-\rho$ and shall thereafter be replaced by the point $(\varphi_0, \varphi_1, h_0, h_1, |t|)$. We denote this deformation by $\Delta_0(P)$. Under Δ_0, each point P of Π is deformed into its restricted projection $\{\varphi_0, \varphi_1\}$ on I^2.

The deformation $\Delta_\psi(P)$. Let ψ be a restricted transformation such that $H(\psi)$ is finite. Then $A(g_k, \psi)$ is finite. Under $\Delta_\psi(P)$, a point P is first deformed under $\Delta_0(P)$ into its restricted projection $\{\varphi_0, \varphi_1\}$. We continue by applying the deformation $\Delta_\psi(\varphi)$ to φ_0 and φ_1, thereby deforming $\{\varphi_0, \varphi_1\}$ into $\{\psi, \psi\}$. The resulting deformation of Π into the point $\{\psi, \psi\}$ will be denoted by $\Delta_\psi(P)$.

Regularity at infinity. Under $\Delta_\psi(P)$, any subset of points of Π on which W is bounded is deformed through a set of points on which W is bounded, while Π is deformed into the point $\{\psi, \psi\}$. The function W is accordingly regular at infinity. Incidentally we see that the 0-dimensional Betti number R_0 (the 0-connectivity) of the subspace of Π on which W is finite is 1, while the remaining Betti numbers are 0.

9. The homotopy theorem.

Let $F(p)$ be a real function defined on a metric space M of points p. A point q at which F is finite is said to be *homotopically ordinary* relative to F if some neighborhood of p relative to the set $F \leqq F(q)$ admits an F-deformation ([9], p. 30) which displaces p. A point at which F is finite which is not homotopically ordinary is termed *homotopically critical.*

If one is concerned with a single contour g, F is taken as the function $A(g, \varphi)$ defined on the space I^* of restricted transformations φ. A transformation ψ of I^* is said to be *differentially critical* if the harmonic surface defined over the unit disc by the boundary vectors $g[\psi(\theta)]$ is minimal. Otherwise ψ is termed *differentially ordinary.* In the case of two contours g_0 and g_1, F is taken as $W(P)$ with P on Π. A point

$$Q = (\psi_0, \psi_1, \sigma_0, \sigma_1, a) \tag{9.1}$$

of Π for which $a > 0$ will be said to be *differentially critical* if the harmonic surface defined on the ring $a \leqq r \leqq 1$ with boundary vectors

$$p_0 = g_0(\psi_0\sigma_0), \qquad p_1 = g_1(\psi_1\sigma_1) \tag{9.2}$$

is minimal. In case $a = 0$, Q is termed *differentially critical* if the two harmonic surfaces defined over the unit disc by boundary vectors $g_k(\psi_k)$ $(k = 0, 1)$ respectively are minimal.

We shall need the following lemma.

LEMMA 9.1. *If Q is differentially critical, then ψ_0 and ψ_1 in Q are strictly increasing; and if further $a > 0$ in Q, Q is non-degenerate.*

If Q is differentially critical and if $a > 0$, Q is non-degenerate as we shall now show. This can be established by reference to (5.6) and (5.7). In particular, when σ_0 and σ_1 in Q are both degenerate, $W_\rho = T_\rho$ at Q. Moreover, the Fourier coefficients $a_0(0)$ and $a_1(0)$ of p_0 and p_1 respectively in (9.2) reduce to vectors defining points of g_0 and g_1. In this case $[a_1(0) - a_0(0)]^2 \neq 0$ in (5.7), so that at Q, $W_\rho = T_\rho > 0$. In case one of the transformations σ_k, say σ_0, is degenerate while the other σ_1 is non-degenerate, we again make use of (5.6) and (5.7). In terms of the Fourier coefficients of p_1 in (9.2) we find that

$$W_\rho \geqq R_\rho = 2\pi \sum_{m=1}^{\infty} \frac{m^2 \rho^{2m-1}}{(1 - \rho^{2m})^2} [a_1^2(m) + b_1^2(m)] > 0 \qquad (\rho > 0)$$

at Q. Hence Q is non-degenerate if $a > 0$.

That ψ_0 and ψ_1 are strictly increasing if $a = 0$ follows as in Radó [12], p. 75.

We return to the case $a > 0$. If ψ_0 were not strictly increasing, $\psi_0\sigma_0$ would not be strictly increasing, and the reasoning of Radó then shows that the ring minimal surface with boundary vectors (9.2) would reduce to a point. This is impossible since p_0 and p_1 in (9.2) do not reduce to points.

The proof of the lemma is complete.

The homotopy theorem is as follows.

THEOREM 9.1. *A point Q at which $W(Q)$ is finite and which is differentially ordinary is homotopically ordinary.*

We consider three cases.

Case I. $a > 0$ in Q, $W_\rho(Q) \neq 0$. Suppose for simplicity that $W_\rho(Q) < 0$. We subject the points P neighboring Q to a deformation in which ρ increases from a, while the other arguments of P remain fixed. For a sufficiently small neighborhood of Q and sufficiently small variation of ρ this is clearly a W-deformation actually displacing Q. The theorem follows in this case.

Case II. $a > 0$, $W_\rho(Q) = 0$. Let h_k $(k = 0, 1)$ be a continuous transformation.. With

$$p_k = g_k(h_k) \qquad (k = 0, 1), \tag{9.3}$$

set

$$D(p_0, p_1, \rho) = B(h_0, h_1, \rho).$$

Referring to Q in (9.1), set $b_k = \psi_k\sigma_k$. The harmonic surface defined on the ring $a \leqq r < 1$ by the boundary vectors (9.3) when $h_k = b_k$ is not minimal by hypothesis. Moreover, $B_\rho = W_\rho = 0$ when $\rho = a$, $h_k = b_k$ and $P = Q$. It follows from Theorem 6.4 of [7] that there exists a real analytic function $\lambda(\alpha)$

with a period 2π such that when $e = 0$, one at least of the following conditions holds:

$$\frac{\partial}{\partial e} B\{b_0, b_1[\alpha + e\lambda(\alpha)], a\} < 0, \tag{9.4}$$

$$\frac{\partial}{\partial e} B\{b_0[\alpha + e\lambda(\alpha)], b_1, a\} < 0. \tag{9.5}$$

We suppose that (9.4) holds. We shall make use of the relation

$$D(p_0, p_1, \rho) = D(p_0) + D(p_1) + \Omega(p_0, p_1, \rho). \tag{9.6}$$

Taken subject to (9.3) we have here a formula for $B(h_0, h_1, \rho)$. It follows from Lemmas 5.3 and 5.4 that for $B(h_0, h_1, \rho)$ bounded, for (h_0, h_1, ρ) sufficiently near (b_0, b_1, a), and for e sufficiently near 0,

$$\frac{\partial}{\partial e} B\{h_0, h_1[\alpha + e\lambda(\alpha)], \rho\} < \text{constant} < 0. \tag{9.7}$$

Designate $\alpha + e\lambda(\alpha)$ by $\beta(\alpha, e)$. We regard $h_1\beta$ as a deformation in which e is the time and in which h_1 is replaced at the time $t = e$ by $h_1\beta$. We wish to define a similar deformation of φ_1 and τ_1 neighboring ψ_1 and σ_1 respectively.

The transformation σ_k $(k = 0, 1)$ in Q is non-degenerate. Otherwise, as we have seen in the proof of Lemma 9.1, $W_\rho(P) > 0$ at Q, contrary to hypothesis in Case II. Hence σ_1 has an inverse σ_1^{-1} as does τ_1 if the distance $[\sigma_1, \tau_1]$ is sufficiently small. Likewise β has a single-valued inverse if $|e|$ is sufficiently small. When β^{-1} and τ_1^{-1} exist, we determine a Möbius transformation T by the conditions

$$T^{-1} = \beta^{-1}\tau_1^{-1} \qquad (\alpha = \alpha_i;\ i = 1, 2, 3).$$

The transformation T reduces to τ_1 when $e = 0$, and depends continuously on (τ_1, α, e) for $|e|$ and $[\sigma_1, \tau_1]$ sufficiently small.

With T so defined, we introduce the transformation

$$\Phi = \varphi_1\tau_1\beta T^{-1}. \tag{9.8}$$

For PQ and $|e|$ sufficiently small, Φ depends continuously on $(\varphi_1, \tau_1, \alpha, e)$, reducing to φ_1 when $e = 0$. Moreover, Φ is a restricted transformation. For when $\alpha = \alpha_i$,

$$\Phi = \varphi_1\tau_1\beta\beta^{-1}\tau_1^{-1} = \varphi_1(\alpha_i) = \alpha_i \qquad (i = 1, 2, 3).$$

Finally we see from (9.8) that

$$\Phi T \equiv \varphi_1\tau_1\beta.$$

Recall that when P is non-degenerate (that is, τ_0 and τ_1 non-degenerate) and $\rho > 0$,

$$W(P) = W(\varphi_0, \varphi_1, \tau_0, \tau_1, \rho) \equiv B(\varphi_0\tau_0, \varphi_1\tau_1, \rho).$$

In particular,

$$W(\varphi_0, \Phi, \tau_0, T, \rho) \equiv B(\varphi_0\tau_0, \varphi_1\tau_1\beta, \rho) \tag{9.9}$$

if PQ is so small that the transformations involved are non-degenerate and $\rho > 0$. This is an identity in e and the arguments of P. It follows from (9.7) as qualified above that

$$\frac{\partial}{\partial e} W(\varphi_0, \Phi, \tau_0, T, \rho) < \text{constant} < 0 \tag{9.10}$$

if PQ and $|e|$ are sufficiently small and $W(P) \leqq W(Q)$.

We now deform P neighboring Q as follows. We hold $(\varphi_0, \tau_0, \rho)$ fast and replace (φ_1, τ_1) by (Φ, T) at the time $t = e$, increasing t from 0. If PQ and t are sufficiently small, this is a continuous deformation, and by virtue of (9.10) a W-deformation. The theorem follows in Case II.

Case III. $a = 0$ *in* Q. The subspace $\rho = 0$ of Π is the topological image of the product $I^* \times I^*$, where I^* is the space of restricted transformations. On this subspace

$$W(\varphi_0, \varphi_1, \tau_0, \tau_1, 0) = A(g_0, \varphi_0) + A(g_1, \varphi_1).$$

If Q is differentially ordinary, one of the representations $g_k(\psi_k)$, say $g_1(\psi_1)$, fails to determine a minimal surface defined over the disc $r \leqq 1$. As shown in [11] in the case of one contour, there will then exist a transformation $\beta(\alpha, e) = \alpha + e\lambda(\alpha)$ of the type used in Case II of this proof such that

$$\frac{\partial}{\partial e} A(g_1, \varphi_1\beta) < \text{constant} < 0 \tag{9.11}$$

for $|e|$ and $|\varphi_1, \psi_1|$ sufficiently small and A bounded.

Let T be a Möbius transformation such that

$$T^{-1} = \beta^{-1} \qquad (\alpha = \alpha_i;\ i = 1, 2, 3),$$

and set

$$\Phi = \varphi_1\beta T^{-1}.$$

The transformation Φ is seen to be restricted. It reduces to φ_1 when $e = 0$. For $|e|$ sufficiently small, Φ depends continuously on e and φ_1. Moreover,

$$\Phi T \equiv \varphi_1\beta. \tag{9.12}$$

We now deform P neighboring Q as follows. We hold $\varphi_0, \tau_0, \tau_1, \rho$ in P fast and replace φ_1 at the time $t = e \geqq 0$ by Φ. Recall that

$$W(\varphi_0, \Phi, \tau_0, \tau_1, \rho) \equiv A(g_0, \varphi_0) + A(g_1, \Phi) + \Omega[g_0(\varphi_0\tau_0), g_1(\Phi\tau_1), \rho]. \tag{9.13}$$

This is an identity in the components of P and e. The e-derivative of the term Ω in (9.13) tends to 0 with ρ uniformly with respect to the remaining components of P and e, for e sufficiently small. Moreover,

$$A(g_1, \Phi) = A(g_1, \varphi_1\beta)$$

by virtue of (9.12) so that the e-derivative of $A(g_1, \Phi)$ satisfies (9.11) as qualified. It follows from (9.13) that the e-derivative of the left member is negative and bounded from 0 for $|e|$ and PQ sufficiently small and $W(P) \leqq W(Q)$.

The theorem follows in Case III.

10. **The functions $H(\varphi_0, \varphi_1)$ and $A^*(\varphi_0, \varphi_1)$.** We set

$$H(\varphi_0, \varphi_1) = H(\varphi_0) + H(\varphi_1),$$

$$A^*(\varphi_0, \varphi_1) = A(g_0, \varphi_0) + A(g_1, \varphi_1).$$

We shall establish theorems for $H(\varphi_0, \varphi_1)$ and $A(\varphi_0, \varphi_1)$ similar to those established for $H(\varphi)$ and $A(\varphi)$ in [11]. On the parallelogram (6.2) we separate off the region ω_e on which

$$0 \leqq \alpha \leqq 2\pi, \qquad -e < \alpha - \beta < e, \tag{10.1}$$

and let

$$H_e(\varphi_0, \varphi_1), \qquad A_e^*(\varphi_0, \varphi_1), \qquad H_e(\varphi)$$

denote integrals with the integrands of $H(\varphi_0, \varphi_1)$, $A^*(\varphi_0, \varphi_1)$, $H(\varphi)$ respectively, but with the domain ω_e instead of the parallelogram (6.2). Theorem 4.2 of [11] is replaced by the following lemma, the proof remaining essentially the same.

LEMMA 10.1. *As e tends to 0, the ratio of $A_e^*(\varphi_0, \varphi_1)$ to $H_e(\varphi_0, \varphi_1)$ tends to 1 uniformly for any subset of Π for which $A^*(\varphi_0, \varphi_1)$ is bounded. Moreover, $H(\varphi_0, \varphi_1)$ is bounded on sets for which $A^*(\varphi_0, \varphi_1)$ is bounded, and conversely.*

The deformation $\Delta_{\psi_0\psi_1}(\varphi_0, \varphi_1)$. Let (ψ_0, ψ_1) be a pair of restricted transformations for which $H(\psi_k)$ $(k = 0, 1)$ is finite. In the deformation $\Delta_{\psi_0\psi_1}$ of a restricted pair (φ_0, φ_1), we understand that φ_k $(k = 0, 1)$ is replaced at the time t $(0 \leqq t \leqq 1)$ by the transformation

$$\varphi_k^t = t\psi_k(\alpha) + (1 - t)\varphi_k(\alpha). \tag{10.2}$$

LEMMA 10.2. *Under the deformation $\Delta_{\psi_0\psi_1}$, the image $(\varphi_0^t, \varphi_1^t)$ of an arbitrary restricted pair varies so that*

$$H_e(\varphi_0^t, \varphi_1^t) \leqq \max [H_e(\varphi_0, \varphi_1), H_e(\psi_0, \psi_1)]. \tag{10.3}$$

We assume that $H(\varphi_0, \varphi_1)$ is finite. Otherwise the lemma is trivial. Upon making use of the definition of $H(\varphi_0^t, \varphi_1^t)$ we find as in the proof of (4.12), [11], that

$$H_e(\varphi_0^t, \varphi_1^t) \leqq t^2 H_e(\psi_0, \psi_1) + 2t(1 - t)[H_e^{\frac{1}{2}}(\varphi_0)H_e^{\frac{1}{2}}(\psi_0) + H_e^{\frac{1}{2}}(\varphi_1)H_e^{\frac{1}{2}}(\psi_1)] + (1 - t)^2 H_e(\varphi_0, \varphi_1).$$

By virtue of the elementary Cauchy inequality ([6], p. 16), the middle term is at most the product of $2t(1 - t)$ and

$$H_e^{\frac{1}{2}}(\varphi_0, \varphi_1)H_e^{\frac{1}{2}}(\psi_0, \psi_1), \tag{10.4}$$

so that

$$H_e(\varphi_0^t, \varphi_1^t) \leqq [tH_e^{\frac{1}{2}}(\psi_0, \psi_1) + (1 - t)H_e^{\frac{1}{2}}(\varphi_0, \varphi_1)]^2,$$

and the lemma follows directly.

We continue with the following lemma, similar to Lemma 5.1 of [11].

LEMMA 10.3. *Under the deformation* $\Delta_{\psi_0\psi_1}$,

$$\frac{d}{dt}H_e(\varphi_0^t, \varphi_1^t) \leqq H_e^{\frac{1}{2}}(\varphi_0, \varphi_1)[H_e^{\frac{1}{2}}(\psi_0, \psi_1) - H_e^{\frac{1}{2}}(\varphi_0, \varphi_1)] \qquad (0 \leqq t \leqq \tfrac{1}{2}), \tag{10.5}$$

whenever

$$H_e(\varphi_0, \varphi_1) > H_e(\psi_0, \psi_1). \tag{10.6}$$

We find as in the proof of Lemma 5.1 of [11] that when $0 \leqq t \leqq \frac{1}{2}$,

$$\frac{d}{dt}H(\varphi_0^t, \varphi_1^t) \leqq 2tH(\psi_0, \psi_1) + 2(1 - 2t)[H^{\frac{1}{2}}(\varphi_0)H^{\frac{1}{2}}(\psi_0) + H^{\frac{1}{2}}(\varphi_1)H^{\frac{1}{2}}(\psi_1)] + 2(t - 1)H(\varphi_0, \varphi_1).$$

Use may now be made of the Cauchy inequality to replace the middle term by the product of $2(1 - 2t)$ and the term (10.4). Upon factoring the resulting right member, we find that when $0 \leqq t \leqq \frac{1}{2}$,

$$\frac{d}{dt}H(\varphi_0^t, \varphi_1^t) \leqq 2[(1 - t)H^{\frac{1}{2}}(\varphi_0, \varphi_1) + tH^{\frac{1}{2}}(\psi_0, \psi_1)][H^{\frac{1}{2}}(\psi_0, \psi_1) - H^{\frac{1}{2}}(\varphi_0, \varphi_1)].$$

A similar relation is valid when the subscript e is added throughout to the function H, and the lemma follows with ease.

The following lemma is similar to Lemma 5.4 of [11] and has a similar proof.

LEMMA 10.4. *If* ψ_0 *and* ψ_1 *are strictly increasing, then under the deformation* $\Delta_{\psi_0\psi_1}$ *for a fixed* e *and bounded* $A^*(\varphi_0, \varphi_1)$, *the* t*-derivative of*

$$A^*(\varphi_0^t, \varphi_1^t) - H_e(\varphi_0^t, \varphi_1^t) \tag{10.7}$$

tends to 0 *uniformly with respect to* t *as* (φ_0, φ_1) *tends to* (ψ_0, ψ_1).

In proving this lemma, one writes the difference (10.7) in the form

$$[A^*(\varphi_0^t, \varphi_1^t) - A_e^*(\varphi_0^t, \varphi_1^t)] + [A_e^*(\varphi_0^t, \varphi_1^t) - H_e(\varphi_0^t, \varphi_1^t)],$$

and proves that the t-derivative of each bracket tends to 0 uniformly with respect to t as (φ_0, φ_1) tends to (ψ_0, ψ_1). In treating the second bracket one follows the proof of Lemma 5.3 of [11] except for the simplification arising from the fact that ψ_0 and ψ_1 are strictly increasing. On this account the case specified by (5.16) in [11] can never arise except in the trivial case where $\alpha = \beta$. See the correction to [11].

11. **Weak upper-reducibility.** We shall be concerned with weak upper-reducibility at a point,

$$Q = (\psi_0, \psi_1, \sigma_0, \sigma_1, a). \tag{11.1}$$

We begin with a deformation $\delta_Q(P)$.

Understanding that $P = (\varphi_0, \varphi_1, \tau_0, \tau_1, \rho)$, we replace φ_k at the time t by

$$\varphi_k^t = t\varphi_k + (1 - t)\psi_k \qquad (k = 0, 1).$$

We see that φ_k is a restricted transformation, depending continuously on φ_k and t. We understand that ψ_k is fixed. When $t = 0$, $\varphi_k^t = \varphi_k$; when $t = 1$, $\varphi_k^t = \psi_k$. Under $\delta_Q(P)$, the variables τ_0, τ_1, ρ in P shall be held fast. Let P^t be the point replacing P at the time t.

We shall investigate $W(P^t)$. Recall that

$$W(P^t) = A(g_0, \varphi_0^t) + A(g_1, \varphi_1^t) + \Omega[g_0(h_0), g_1(h_1), \rho], \tag{11.2}$$

where

$$h_k = t\varphi_k\tau_k + (1 - t)\psi_k\tau_k \qquad (k = 0, 1).$$

Observe that

$$\frac{dh_k}{dt} = \varphi_k\tau_k - \psi_k\tau_k,$$

so that this derivative tends to 0 as P tends to Q. It follows from Lemma 5.5 that the t-derivative of the term Ω in (11.2) exists and tends to 0 uniformly as PQ tends to 0, provided ρ in P is bounded from 1.

The following lemma is a restricted counterpart of Lemma 5.4 of [11].

LEMMA 11.1. *If ψ_0 and ψ_1 in Q are strictly increasing, then under $\delta_Q(P)$ for a fixed e and bounded $W(P)$ the t-derivative of*

$$W(P^t) - H_e(\varphi_0^t, \varphi_1^t) \tag{11.3}$$

tends to 0 uniformly with respect to t as P tends to Q.

By virtue of the definition of $W(P)$, the difference (11.3) equals (cf. §10)

$$A^*(\varphi_0^t, \varphi_1^t) - H_e(\varphi_0^t, \varphi_1^t) + \Omega,$$

where Ω is the last term in $W(P^t)$. The lemma follows at once from Lemma 10.4 and from our remarks concerning Ω.

The fundamental theorem of this section is as follows.

THEOREM 11.1. *The function $W(P)$ is weakly upper-reducible at each point Q of Π at which $W(Q)$ is finite.*

We give the proof under two cases.

Case I. *Q is differentially critical.* In this case ψ_0 and ψ_1 in Q are strictly increasing by virtue of Lemma 9.1. Hence we can apply Lemma 10.4 as well as the other lemmas of §10. The theorem follows in Case I from the lemmas of §10 and Lemma 11.1, as did Theorem 5.1 of [11] from the corresponding lemmas of §5 of [11].

Case II. *Q is differentially ordinary.* The deformations used in proving the homotopy theorem in §10 suffice to prove the theorem in this case.

12. **The existence of minimal surfaces of non-minimum type.** We have now shown that $W(P)$ is boundedly compact on Π, regular at infinity, and weakly upper-reducible. We have seen that a homotopic critical point Q is a differential critical point and determines a ring minimal surface bounded by g_0 and g_1 if $\rho > 0$ in Q, and two disc type minimal surfaces bounded by g_0 and g_1 respectively if $\rho = 0$ in Q. We have further shown that the subspace of Π on which $W(P)$ is finite has a Betti number $R_0 = 1$ with its other Betti numbers null. *The general theory as developed in* [8] *and* [9] *thus applies.*

We recall certain definitions. See [9], p. 39.

Let $F(p)$ be a function defined on a metric space M of points p. By a *homotopic critical set* σ at the level c is meant a set of homotopic critical points on which $F = c$ and which is at a positive distance from other homotopic critical points at the level c. The set σ is closed relative to the set at the level c. A homotopic critical set σ at which F assumes a value less than at points of $M - \sigma$ sufficiently near σ will be called a *minimizing* set. A point q of M will be termed *non-minimizing* if in every neighborhood of q there is a point p such that $F(p) < F(q)$. A homotopic critical set which is not a minimizing set contains at least one non-minimizing point of F.

These terms will be carried over to the space Π and function W. We shall also apply the terms minimizing set or non-minimizing point to minimal surfaces whenever the terms apply to the corresponding points of Π. We extend these terms to the function $A(g_k, \varphi)$.

Let S_k $(k = 0, 1)$ be a minimizing set of disc minimal surfaces belonging to the function $A(g_k, \varphi)$, and let H_k be the set of points on all surfaces of S_k. If H_0 and H_1 do not intersect, S_0 and S_1 will be said to be *spatially disjoint.* The transformations φ_k representing the surfaces of S_k form a compact set ω_k. Hence H_0 is bounded from H_1 if S_0 and S_1 are spatially disjoint.

To the hypotheses (6a) and (6b) (alternately (1a)), we add the following hypothesis.

(12a) *Corresponding to the curves* g_k $(k = 0, 1)$, *there shall exist spatially disjoint minimizing sets* S_k *of disc minimal surfaces.*

This condition is satisfied if g_0 and g_1 are separated by an $(n - 1)$-plane.

It will be well to review the nature of the previous hypotheses:

(1a) *This is the chord arc hypothesis.*

(6a) *This concerns the representation of* g_0 *and* g_1, *and postulates rectifiability.*

(6b) *This imposes a Lipschitz condition on* g_0' *and* g_1'. *It will eventually be replaced by* (1a).

The minimizing set Ω. We suppose hypothesis (12a) is satisfied. In the subspace $\rho = 0$ of Π, let Ω be the subset of points $\{\psi_0, \psi_1\}$ determined by the product $\omega_0 \times \omega_1$. Let h be a continuous or degenerate transformation, and let $E^k(h)$ be the disc harmonic surface with boundary vector $g_k(h)$ $(k = 0, 1)$. If $\bar{N}$ is the closure of a sufficiently small neighborhood of Ω of points $\{\varphi_0, \varphi_1\}$, no surface $E^0(\varphi_0)$ will intersect a surface $E^1(\varphi_1)$ for $\{\varphi_0, \varphi_1\}$ on $\bar{N}$. If τ_k is an arbitrary

Möbius or degenerate transformation, the surface $E^k(\varphi_k\tau_k)$ as a point set is on $E^k(\varphi_k)$. Hence $E^0(\varphi_0\tau_0)$ and $E^1(\varphi_1\tau_1)$ are bounded, one from the other, for $\{\varphi_0, \varphi_1\}$ on N.

For $\{\varphi_0, \varphi_1\}$ on N, the Fourier coefficients $a_k(0)$ of $g_k(\varphi_k\tau_k)$ $(k = 0, 1)$ are bounded from one another since they represent points on $E^0(\varphi_0)$ and $E^1(\varphi_1)$ respectively; hence

$$| a_1(0) - a_0(0) | \tag{12.1}$$

in (5.7) is bounded from 0. It follows from (5.5), (5.6) and (5.7) that Ω is a minimizing set for $W(P)$.

We can now prove a special theorem.

THEOREM 12.1. *If Hypotheses* (6a), (6b) *and* (12a) *hold, and if g_0 and g_1 bound a ring minimal surface Σ belonging to a minimizing set, there either exists a ring minimal surface of non-minimizing type bounded by g_0 and g_1, or else a disc minimal surface of non-minimizing type bounded by g_0 or by g_1.*

Let Q be the point determining Σ, and let ω be a minimizing set of $W(P)$ to which Q belongs. The set ω may intersect Ω. But $\rho = 0$ on Ω and $\rho > 0$ at Q. There is accordingly a subset of ω, say $\bar{\omega}$, which is a minimizing critical set, which contains Q, and which does not intersect Ω.

It follows from the general theory that there exists a homotopic critical set σ whose first type number is at least 1. Such a set is not a minimizing set, and accordingly contains at least one homotopic critical point Q^* of non-minimizing type. If $\rho > 0$ in Q^*, Q^* determines a ring minimal surface in accordance with the homotopy theorem. If $\rho = 0$ in Q^*, Q^* determines two disc minimal surfaces, at least one of which must be of non-minimizing type.

The proof of the theorem is complete.

The hypotheses of the theorem permit a particularly simple specialization as follows. If g_0 and g_1 are separated by an $(n - 1)$-plane, (12a) is satisfied. If g_0 and g_1 each possess a convex $(n - 1)$-plane projection, g_0 and g_1 each bound a minimizing disc minimal surface, and no other disc minimal surface. Hence we have the following corollary of the theorem.

COROLLARY. *If g_0 and g_1 bound a ring minimal surface belonging to a minimizing set, if g_0 and g_1 are separated by an $(n - 1)$-plane and possess convex plane projections, there exists a ring minimal surface of non-minimum type bounded by g_0 and g_1.*

The subspace $\rho = 0$. If a point Q for which $\rho = 0$ is a homotopic critical point of $W(P)$, the restricted projection $\{\psi_0, \psi_1\}$ of Q on the subspace $\rho = 0$ of Π is a homotopic critical point of

$$A(g_0, \varphi_0) + A(g_1, \varphi_1), \tag{12.2}$$

while ψ_k is a homotopic critical point of $A(g_k, \varphi_k)$ $(k = 0, 1)$. One is here concerned with a product space $X \times Y$ and a function

$$A(x) + B(y), \tag{12.3}$$

in which $A(x)$ is defined on X, and $B(y)$ on Y.

The following question immediately arises. Is the r-th type number of a

critical set of the function (12.3) determined by the i-th and j-th type numbers of the corresponding critical sets of $A(x)$ and $B(y)$ respectively for $i + j = r$? A. E. Pitcher [14] has investigated this question and given an answer under certain general conditions. The result of Pitcher is easily verified if X and Y are Euclidean spaces with Cartesian coördinates (x) and (y) respectively and if $A(x)$ and $B(y)$ are non-degenerate quadratic forms. The theorem in the general case appears rather deep and involves new functional and topological considerations.

Reduction of hypotheses. Theorem 12.1 and its corollary have been established under conditions (6a), (6b), and (12a). Condition (6b) can be replaced by the much weaker chord arc condition of the introduction, leading to the following result.

THEOREM 12.2. *Theorem* 12.1 *and its corollary hold if condition* (6b) *on* g_0 *and* g_1 *is replaced by the chord arc condition on* g_0 *and* g_1.

The details of the proof of this theorem are similar to the details of the proof developed in [12] for the case of one contour.

BIBLIOGRAPHY

1. C. CARATHÉODORY, *Conformal Representation*, Cambridge University Press, 1932.
2. R. COURANT, *The existence of minimal surfaces of given topological structure under prescribed boundary conditions*, Acta Mathematica, vol. 72(1940), pp. 51-97.
3. R. COURANT, *Plateau's problem and Dirichlet's principle*, Annals of Mathematics, vol. 38(1937), pp. 679-724.
4. J. DOUGLAS, *The problem of Plateau for two contours*, Journal of Mathematics and Physics, vol. 10(1931), pp. 315-359.
5. J. DOUGLAS, *Minimal surfaces of higher topological structure*, Annals of Mathematics, vol. 40(1939), pp. 205-298.
6. G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, *Inequalities*, Cambridge University Press, 1934.
7. MARSTON MORSE, *The first variation in minimal surface theory*, this Journal, vol. 6(1940), pp. 263-289.
8. MARSTON MORSE, *Rank and span in functional topology*, Annals of Mathematics, vol. 41(1940), pp. 419-454.
9. MARSTON MORSE, *Functional Topology and Abstract Variational Theory*, Mémorial des Sciences Mathématiques, Paris, 1939.
10. MARSTON MORSE AND C. TOMPKINS, *Minimal surfaces of unstable type*, Bulletin of the American Mathematical Society, vol. 46(1940), p. 222.
11. MARSTON MORSE AND C. TOMPKINS, *The existence of minimal surfaces of general critical types*, Annals of Mathematics, vol. 40(1939), pp. 443-472. (Corrections for this paper appear in the Annals of Mathematics, vol. 42(1941), p. 331.
12. MARSTON MORSE AND C. TOMPKINS, *Minimal surfaces of unstable type by a new mode of approximation*, Annals of Mathematics, vol. 42(1941), pp. 62-72.
13. MARSTON MORSE AND C. TOMPKINS, *Unstable minimal surfaces of higher topological structure*, Proceedings of the National Academy of Sciences, vol. 26(1940), pp. 713-716.
14. EVERETT PITCHER, *The cap heights of a sum function on a product space*, Bulletin of the American Mathematical Society, vol. 46(1940), p. 887.
15. TIBOR RADÓ, *On the problem of Plateau*, Ergebnisse der Mathematik, vol. 2(1933).
16. MAX SHIFFMAN, *The Plateau problem for non-relative minima*, Annals of Mathematics, vol. 40(1939), pp. 834-854.

THE INSTITUTE FOR ADVANCED STUDY AND PRINCETON UNIVERSITY.

THE BEST POLYNOMIAL APPROXIMATION OF FUNCTIONS POSSESSING DERIVATIVES

By J. Shohat

1. **Best polynomial approximation and interpolation.** Let $f(x)$ be continuous in the finite closed interval (a, b). Denote by $\Pi_n(x; f)$ the polynomial, of degree not exceeding n, of best approximation, in the sense of Tchebycheff, to $f(x)$ on (a, b). Its well-known characteristic property ([6], p. 76)[1] is the following

Property E. Among the points in (a, b) where $f(x) - \Pi_n(x; f)$ attains its extreme values $\pm E_n(f)$ ("points of deviation") there exist at least $n + 2$ points, $x_{1,n} < x_{2,n} < \cdots < x_{n+2,n}$, where the signs of $f(x_{i,n}) - \Pi_n(x_{i,n}; f)$ alternate.

It follows that $f(x) - \Pi_n(x; f)$ vanishes at $n + 1$ points, at least, in (a, b). Thus, *every function $f(x)$, continuous in (a, b), possesses therein a uniquely determined sequence of Lagrange interpolation polynomials (LIP) $\Pi_n(x; f)$ $(n = 1, 2, \cdots)$, characterized by the property that $\lim_{n \to \infty} \Pi_n(x; f) = f(x)$ holds uniformly on (a, b), and more rapidly than for any other sequence of polynomials.*

It is this close connection between the theory of interpolation and that of best approximation that is being utilized in the present paper. In this way we establish in a simple manner certain general theorems on best approximation for functions possessing derivatives—as many as needed in the discussion. Some of these results have been stated previously in a more restricted form and proved more elaborately by S. Bernstein [1, 2] (cf. also [5]). In our more general statements we replace, among other things, the sign $>$ by $\geqq$. The usefulness of such extension is seen, for example, from the fact that it enables us to draw at once, *without any further consideration*, the conclusion that

$$2\left(\frac{b-a}{4}\right)^{n+1} \cos (n+1) \operatorname{arc} \cos \left(\frac{2x-a-b}{b-a}\right)$$

deviates the least from zero on (a, b) among all polynomials of the form $x^n + \cdots$, a conclusion which is often used in the general theory of polynomials.

Our chief object is to establish new results dealing with the distribution of the points of deviation; we also treat briefly the highest coefficients of $\Pi_n(x; f)$.

Our main and practically only tool is the following elementary relation in the theory of interpolation:

$$f(x) = L_n(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=1}^{n+1} (x - c_i) \tag{1}$$

$$(a \leqq x \leqq b;\ a \leqq c_1 < c_2 < \cdots < c_{n+1} \leqq b),$$

Received February 10, 1941; presented to the American Mathematical Society, May 3, 1941.

[1] Numbers in brackets refer to the bibliography at the end of the paper.

where $L_n(x)$, of degree not exceeding n, is a LIP for $f(x)$: $L_n(c_i) = f(c_i)$ $(i = 1, 2, \cdots, n + 1)$. ($\xi$ here and hereafter denotes a certain quantity lying *inside* (a, b), different in different formulas.)

2. Some direct consequences of Property E.

(i) *$f(x) - \Pi_n(x; f)$ has $n + 2 + m$ points of deviation, $m > 0$, where $f(x) - \Pi_n(x; f)$ alternates sign, if and only if*

$$E_n(f) = E_{n+1}(f) = \cdots = E_{n+m}(f), \tag{2}$$

which is equivalent to

$$\Pi_n(x; f) \equiv \Pi_{n+1}(x; f) \equiv \cdots \equiv \Pi_{n+m}(x; f). \tag{3}$$

(ii) Assuming the existence of $f^{(n+1)}(x)$ in (a, b), we write, by virtue of (1), the following relation:

$$f(x) - \Pi_n(x; f) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=1}^{n+1} (x - \xi_{i,n}) \tag{4}$$

$$(a \leqq x \leqq b;\ x_{i,n} < \xi_{i,n} < x_{i+1,n}),$$

where the $x_{i,n}$ are the points of deviation characterized by Property E. Assume further that $f^{(\nu)}(x)$ *does not change sign in* (a, b),[2] say (no loss of generality),

$$f^{(\nu)}(x) \geqq 0, \qquad a \leqq x \leqq b. \tag{5}$$

It follows that $f(x) - \Pi_n(x; f)$, $n < \nu$, has in (a, b) at most $\nu - 1$ interior points of deviation; hence, the total number of points of deviation does not exceed $\nu + 1$. This is a trivial application of Rolle's theorem whose complete statement is as follows: $F(a) = F(b) = 0$ implies that $F'(c)$ *vanishes and changes sign*, $a < c < b$ (under well-known conditions on $F(x)$, $F'(x)$). In particular, *if $f^{(n+1)}(x) \geqq 0$ in (a, b), then $f(x) - \Pi_n(x; f)$ has therein precisely $n + 2$ points of deviation $x_{i,n}$* :

$$a = x_{1,n} < x_{2,n} < \cdots < x_{n+1,n} < x_{n+2,n} = b. \tag{6}$$

Moreover, (4) *yields the important equalities*

$$f(x_{i,n}) - \Pi_n(x_{i,n}; f) = (-1)^{n+i} E_n(f) \qquad (i = 1, 2, \cdots, n + 2). \tag{7}$$

(iii) Combining (i) and (ii), we conclude that *if neither $f^{(n)}(x)$ nor $f^{(n+1)}(x)$ changes sign in (a, b), then $E_{n-1}(f) \neq E_n(f)$, and this is equivalent to* $\Pi_{n-1}(x; f) \not\equiv \Pi_n(x; f)$. The case of $a = -1$, $b = 1$, $f(x) = x^{n+1}$ (see below) shows that the last conclusion does not always hold, if one at least of the conditions stated is not satisfied.

[2] This means here and hereafter either $\geqq 0$ or $\leqq 0$, but not $\equiv 0$.

3. **Bounds for $E_n(f)$.**

THEOREM I. *If in* (a, b) (i) $|f^{(n+1)}(x)| \leqq M_{n+1}$, *or* (ii) $f^{(n+1)}(x) \geqq m > 0$, *then correspondingly*

$$\text{(i)} \qquad E_n(f) \leqq 2\left(\frac{b-a}{4}\right)^{n+1} \cdot \frac{M_{n+1}}{(n+1)!},$$

or

$$\text{(ii)} \qquad E_n(f) \geqq 2\left(\frac{b-a}{4}\right)^{n+1} \cdot \frac{m_{n+1}}{(n+1)!}.$$

Consider $L_n(x)$, a LIP, coinciding with $f(x)$ at the $n + 1$ zeros of the trigonometric polynomial

$$T_{n+1}(x) = 2\left(\frac{b-a}{4}\right)^{n+1} \cos (n+1)\theta = x^{n+1} + \cdots \qquad \left(\cos\theta = \frac{2x-a-b}{b-a}\right), \tag{8}$$

and apply (1):

$$f(x) - L_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \cdot 2\left(\frac{b-a}{4}\right)^{n+1} \cos (n+1)\theta \qquad (a \leqq x \leqq b). \tag{9}$$

Assumption (i) leads at once to

$$E_n(f) \leqq \max_{a \leqq x \leqq b} |f(x) - L_n(x)| \leqq 2\left(\frac{b-a}{4}\right)^{n+1} \cdot \frac{M_{n+1}}{(n+1)!}.$$

Under assumption (ii), (9) shows that at the $n + 2$ points in (a, b) where $\cos (n+1)\theta = \pm 1$, $f(x) - L_n(x)$ takes values of alternating signs which numerically are not less than $2[\frac{1}{4}(b-a)]^{n+1} m_{n+1}/(n+1)!$, and (ii) follows, by the fundamental result of de la Vallée-Poussin which gives a lower bound for $E_n(f)$ ([6], p. 78).

COROLLARY 1.

$$0 < m_{n+1} \leqq f^{(n+1)}(x) \leqq M_{n+1} \quad \textit{in } (a, b) \textit{ implies}$$

$$2\left(\frac{b-a}{4}\right)^{n+1} \cdot \frac{m_{n+1}}{(n+1)!} \leqq E_n(f) \leqq 2\left(\frac{b-a}{4}\right)^{n+1} \cdot \frac{M_{n+1}}{(n+1)!}. \tag{10}$$

Note that (10) gives the "best possible results". In fact, it yields $E_n(x^{n+1}) = 2[\frac{1}{4}(b-a)]^{n+1}$, whence, by (9),

$$\Pi_n(x; x^{n+1}) = x^{n+1} - 2\left(\frac{b-a}{4}\right)^{n+1} \cos (n+1) \arccos \left(\frac{2x-b-a}{b-a}\right).$$

COROLLARY 2. *If in* (a, b) $0 < m_{n+1} \leqq f^{(n+1)}(x) \leqq M_{n+1}$, $0 < m_{n+2} \leqq f^{(n+2)}(x) \leqq M_{n+2}$, *then*

$$4\,\frac{n+2}{b-a}\,\frac{m_{n+1}}{M_{n+2}} \leqq \frac{E_n(f)}{E_{n+1}(f)} \leqq 4\,\frac{n+2}{b-a}\,\frac{M_{n+1}}{m_{n+2}},$$

whence

$$M_{n+1} > \frac{b-a}{4(n+2)} m_{n+2}.$$

Write (10) for $E_n(f)$ and $E_{n+1}(f)$, and observe that here $E_n(f) > E_{n+1}(f)$.

4. **Comparison of $E_n(f)$ and $E_n(\varphi)$.**

THEOREM II. *$f^{(n+1)}(x) \geqq \varphi^{(n+1)}(x) \geqq 0$ in (a, b) implies $E_n(f) \geqq E_n(\varphi)$.*

Apply (6) and (7) to $f(x)$ and $\varphi(x)$. Let $z_1 = a < z_2 < \cdots < z_{n+2} = b$ denote the points of deviation for $\varphi(x) - \Pi_n(x; \varphi)$. Consider the function

$$\Delta(x) = f(x) - \Pi_n(x; f) - [\varphi(x) - \Pi_n(x; \varphi)].$$

We have, on the one hand, by virtue of (6) and (7),

$$(11) \quad (-1)^{n+1}\Delta(a) = (-1)^{n+1}[(-1)^{n+1}E_n(f) - (-1)^{n+1}E_n(\varphi)] = E_n(f) - E_n(\varphi).$$

On the other hand, if $E_n(f) < E_n(\varphi)$, then at the points $x = z_i$ $(i = 1, 2, \cdots, n + 2)$ $\Delta(x)$ has the sign of $\Pi_n(x; \varphi) - \varphi(x)$; that is, $(-1)^{n+i-1}\Delta(z_i) > 0$, whence, by (1),

$$(12) \quad \Delta(x) = \frac{F^{(n+1)}(\xi)}{(n+1)!} \prod_{i=1}^{n+1} (x - \eta_i), \qquad a \leqq x \leqq b$$
$$(F(x) = f(x) - \varphi(x);\, z_i < \eta_i < z_{i+1}).$$

Hence,

$$(-1)^{n+1}\Delta(a) \geqq 0$$

since, by assumption, $F^{(n+1)}(x) \geqq 0$ in (a, b), while (11) gives $(-1)^{n+1}\Delta(a) < 0$ if $E_n(f) < E_n(\varphi)$.

5. **Distribution of the points of deviation.** Denote by δ_n the minimum distance between two successive points of deviation $x_{i,n}$ $(i = 1, 2, \cdots, n + 2)$ for $f(x) - \Pi_n(x; f)$, where this difference alternates sign. It follows immediately from (4) that[3]

$$(13) \quad |f^{(n+1)}(x)| \leqq M_{n+1} \text{ in } (a, b) \text{ implies } \delta_n \geqq \frac{E_n \cdot (n+1)!}{M_{n+1} \cdot (b-a)^n}$$

and that

$$(14) \quad 0 < m_{n+1} \leqq f^{(n+1)}(x) \leqq M_{n+1} \text{ in } (a, b) \text{ implies } \delta_n \geqq \frac{m_{n+1}}{M_{n+1}} \frac{b-a}{2^{2n+1}}.$$

[3] de la Vallée-Poussin ([6], pp. 81–83), by means of an elaborate analysis, puts a factor 2 in the numerator.

This estimate (as well as that of de la Vallée-Poussin[3]) is obviously very rough. Far more important is the following

THEOREM III. *Assume neither of the functions*

$$\varphi^{(n)}(x),\ f^{(n+1)}(x),\ \frac{f^{(n+1)}(x)}{E_n(f)} \mp \frac{\varphi^{(n+1)}(x)}{E_{n-1}(\varphi)}$$

changes sign in (a, b). *Let* $\{x_{i,n}\}$, $\{z_{j,n-1}\}$ $(i = 1, 2, \cdots, n+2; j = 1, 2, \cdots, n+1)$ *denote the points of deviation for* $f(x) - \Pi_n(x; f)$ *and* $\varphi(x) - \Pi_{n-1}(x; \varphi)$ *respectively. Then* (see (6), (7))

$$\begin{aligned} a = x_{1,n} = z_{1,n-1} < x_{2,n} < z_{2,n-1} \\ < \cdots < z_{n,n-1} < x_{n+1,n} < z_{n+1,n-1} = x_{n+2,n} = b. \end{aligned} \tag{15}$$

Introduce the functions

$$F_{1,2}(x) = f(x) - \Pi_n(x;f) \mp \frac{E_n(f)}{E_{n-1}(\varphi)}[\varphi(x) - \Pi_{n-1}(x;\varphi)].$$

By our assumptions, $F_\sigma^{(n+1)}(x)$ $(\sigma = 1, 2)$ do not change sign in (a, b), so that $F_\sigma(x)$ may have therein at most $n + 1$ zeros. To abbreviate, write x_i, z_j for $x_{i,n}$, $z_{j,n}$. It is seen that $F_\sigma(x_i)F_\sigma(x_{i+1}) < 0$, $F_\sigma(z_j)F_\sigma(z_{j+1}) < 0$, unless $x_i = z_j$ for a certain pair (i, j); in this case one (but not both) of $F_1(x)$, $F_2(x)$ is zero at x_i, and x_i, assumed $\neq a, b$, is a multiple root. (Since $f(x) - \Pi_n(x;f)$ and $\varphi(x) - \Pi_{n-1}(x;\varphi)$ both have horizontal tangents at $x_i = z_j$, $F_1(x)$ or $F_2(x)$ has at least a double root at x_i.) Assume now Theorem III does not hold. Then several z_j may fall inside a certain (x_i, x_{i+1}); also some of the z_j may coincide with some of the x_i. Compare the number of variations of sign in the sequences $\{f(x_k) - \Pi_n(x_k;f)\}$ $(k = 1, 2, \cdots, n+2)$, $\{F_\sigma(x)\}$, in the latter two sequences x taking the values x_i, z_j, arranged in increasing order of magnitude. It is readily seen that if two or more z_j lie between x_i and x_{i+1}, the number of variations of sign in $(x_i, \cdots, x_{i+1})$, for at least one sequence $\{F_\sigma(x)\}$, is greater than for the sequence $\{f(x_k) - \Pi_n(x_k;f)\}$. Furthermore, if $x_i = z_j$ $(\neq a, b)$ and, say, $F_1(x_i) = 0$, while $F_1(x_{i-1})F_1(x_{i+1}) \neq 0$, the number of zeros of $F_1(x)$ in (x_{i-1}, x_{i+1}) is at least 2, which is precisely the number of variations of sign for $\{f(x_k) - \Pi_n(x_k;f)\}$ in the same interval, since x_i is a multiple root for $F_1(x)$. For the same reason, if $F_1(x_i) = F_1(x_{i+1}) = \cdots = F_1(x_{i+l}) = 0$ $(l \geqq 1)$, the number of zeros of $F_1(x)$ in (x_{i-1}, x_{i+l+1}) is greater than the number of variations of sign for $\{f(x_k) - \Pi_n(x_k;f)\}$ in the same interval, whether $(x_i - a)(x_{i+l} - b) \neq 0$ or $= 0$ (with slight proper modifications in the latter case). In a similar manner we treat the case where the z_j are so located that no two (or more) z_j lie between two successive x_i's, and one or more z_j coincide with one or more x_i, but no two (or more) *successive* x_i's coincide with the z_j's.[4]

[4] This case was pointed out to the writer by the referee. For the underlying ideas in the above reasoning consult the more elaborate analysis of S. Bernstein ([2], pp. 85-87), who proves the particular statement of Corollary 1 to Theorem III only.

Thus, rejecting the validity of Theorem III leads to the conclusion that at least one $F_\sigma(x)$ has in (a, b) more then $n + 1$ zeros, and this is impossible. Taking $\varphi(x) = x^n$, we derive

COROLLARY 1. *If $f^{(n+1)}(x)$ does not change sign in (a, b), then*

$$
\begin{aligned}
&a = x_{1,n} < x_{2,n} < c_1 < x_{3,n} < c_2 < \cdots < c_{n-1} < x_{n+1,n} < x_{n+2,n} = b, \\
&c_i = \frac{b - a}{2} \cos \frac{(n - i)\pi}{n} + \frac{b + a}{2} \qquad (i = 1, 2, \cdots, n - 1).
\end{aligned} \tag{16}
$$

Remarks. (i) If each of the two functions

$$\frac{f^{(n+1)}(x)}{E_n(f)} - \frac{\varphi^{(n+1)}(x)}{E_{n-1}(\varphi)}, \qquad \frac{f^{(n+1)}(x)}{E_n(f)} + \frac{\varphi^{(n+1)}(x)}{E_{n-1}(\varphi)}$$

keeps a constant sign in (a, b), then adding or subtracting, according as these signs are alike or opposite, we find correspondingly that $f^{(n+1)}(x)$ or $\varphi^{(n+1)}(x)$ keeps a constant sign in (a, b). (ii) If $\varphi^{(n+1)}(x)$ keeps a constant sign, say, $\geqq 0$ in (a, b) (including the case of $\varphi^{(n+1)}(x) \equiv 0$) and if of the two functions

$$\frac{f^{(n+1)}(x)}{E_n(f)} - \frac{\varphi^{(n+1)}(x)}{E_{n-1}(\varphi)}, \qquad \frac{f^{(n+1)}(x)}{E_n(f)} + \frac{\varphi^{(n+1)}(x)}{E_{n-1}(\varphi)}$$

the first $\geqq 0$ or the second $\leqq 0$ in (a, b), then the other function and $f^{(n+1)}(x)$ both keep the same and constant sign in (a, b), as is seen directly. Hence, we have the following conclusion. *Theorem* III *holds if* the conditions stated therein are modified as follows: *neither $\varphi^{(n)}(x)$ nor $\varphi^{(n+1)}(x)$ changes sign in (a, b), say, $\varphi^{(n+1)}(x) \geqq 0$ (including the case of $\varphi^{(n+1)}(x) \equiv 0$), and of the two functions*

$$\frac{f^{(n+1)}(x)}{E_n(f)} - \frac{\varphi^{(n+1)}(x)}{E_{n-1}(\varphi)}, \qquad \frac{f^{(n+1)}(x)}{E_n(f)} + \frac{\varphi^{(n+1)}(x)}{E_{n-1}(\varphi)}$$

the first $\geqq 0$ or the second $\leqq 0$ in (a, b). Moreover, under these modified conditions, neither of the functions $f^{(n+1)}(x)/E_n(f) \mp \varphi^{(n+1)}(x)/E_{n-1}(\varphi)$ vanishes identically in (a, b) (disregarding the trivial case where $f^{(n+1)}(x) \equiv 0, \varphi^{(n+1)}(x) \equiv 0$). For then accordingly,

$$f(x) = \pm \frac{E_n(f)}{E_{n-1}(\varphi)} \varphi(x) + P(x),$$

where $P(x)$ is a polynomial of degree not exceeding n, whence

$$E_n(f) = \frac{E_n(f)}{E_{n-1}(\varphi)} \cdot E_n(\varphi), \qquad E_n(\varphi) = E_{n-1}(\varphi).$$

This is incompatible with the condition that neither $\varphi^{(n)}(x)$ nor $\varphi^{(n+1)}(x)$ changes sign in (a, b) (cf. (iii) in §2).

Taking $\varphi(x) \equiv f(x)$, we derive

COROLLARY 2. *If neither $f^{(n)}(x)$ nor $f^{(n+1)}(x)$ changes sign in (a, b), then the points of deviation $\{x_{i,n}\}$ and $\{x_{j,n-1}\}$ alternate; namely,*

$$\begin{aligned} a = x_{1,n} = x_{1,n-1} < x_{2,n} < x_{2,n-1} < x_{3,n} \\ < \cdots < x_{n+1,n} < x_{n+1,n-1} = x_{n+2,n} = b. \end{aligned} \tag{17}$$

Hence, if each $f^{(p)}(x)$, $p \geqq N$, keeps a constant sign in (a, b), then (17) holds for all $n \geqq N$.

In fact, here $E_{n-1}(f) > E_n(f)$, and

$$\frac{f^{(n+1)}(x)}{E_n(f)} \mp \frac{f^{(n+1)}(x)}{E_{n-1}(f)} = \frac{1}{E_n(f)} f^{(n+1)}(x).$$

$\{1 \mp E_n(f)/E_{n-1}(f)\}$ does not change sign in (a, b).

6. **Further generalizations of Theorem** III.

THEOREM IV. *If neither $f^{(n+1)}(x)$ nor $\varphi^{(n+1)}(x)$ changes sign in (a, b), then $E_n(f)/E_n(\varphi) = |f^{(n+1)}(\xi)/\varphi^{(n+1)}(\xi)|$ at a certain point ξ inside (a, b).*[5]

Consider the functions

$$\psi_{1,2}(x) = \frac{f^{(n+1)}(x)}{E_n(f)} \mp \frac{\varphi^{(n+1)}(x)}{E_n(\varphi)}.$$

If neither function vanishes inside (a, b), we introduce, as above in proving Theorem III,

$$f_{1,2}(x) = f(x) - \Pi_n(x; f) \mp \frac{E_n(f)}{E_n(\varphi)} [\varphi(x) - \Pi_n(x; \varphi)],$$

and employ the same reasoning. This leads to inequalities analogous to (15), namely,

$$\begin{aligned} a = x_{1,n} = z_{1,n} < x_{2,n} < z_{2,n} < x_{3,n} \\ < \cdots < x_{n+1,n} < z_{n+1,n} < x_{n+2,n} = z_{n+2,n} = b, \end{aligned}$$

and, interchanging $f(x)$ and $\varphi(x)$, we get precisely the same inequalities with the $z_{i,n}$ and $x_{i,n}$ interchanged. This is manifestly absurd. The alternative assumption $z_{i,n} = x_{i,n}$, for $i = 1, 2, \cdots, n + 2$, must be rejected; for it yields, since we may assume $f^{(n+1)}(x) \geqq 0$, $\varphi^{(n+1)}(x) \geqq 0$ in (a, b), $f_1(x_{i,n}) = 0$ $(i = 1, 2, \cdots, n + 2)$, and this is impossible. Thus, at least one of $\psi_{1,2}(x)$ vanishes inside (a, b), and our statement follows. Theorem IV gives at once the statement of Theorem II, also the following one: $f^{(n+1)}(x) > \varphi^{(n+1)}(x) > 0$ *inside* (a, b) *implies* $E_n(f) > E_n(\varphi)$.

[5] This holds, in a trivial sense, if one or both derivatives $f^{(n+1)}(x)$, $\varphi^{(n+1)}(x)$ vanish identically.

COROLLARY 1. *If neither* $f^{(n+1)}(x)$ *nor* $\varphi^{(n+1)}(x)$ *changes sign in* (a, b), *then*

$$\min_{a \leq x \leq b} \left| \frac{f^{(n+1)}(x)}{\varphi^{(n+1)}(x)} \right| \leq \frac{E_n(f)}{E_n(\varphi)} \leq \max_{a \leq x \leq b} \left| \frac{f^{(n+1)}(x)}{\varphi^{(n+1)}(x)} \right|. \tag{18}$$

Taking $\varphi(x) = x^{n+1}$, we derive

COROLLARY 2. *If* $f^{(n+1)}(x)$ *does not change sign in* (a, b), *then*

$$E_n(f) = 2\left(\frac{b-a}{4}\right)^{n+1} \cdot \frac{|f^{(n+1)}(\xi)|}{(n+1)!}. \tag{19}$$

From (19) relation (10) follows directly. (19) shows the advantage of the best approximation over that furnished by Taylor's formula. It further yields the following property of indefinitely differentiable functions. *If* $f(x)$ *has in* (a, b) *all derivatives and if therein* $f^{(n)}(x) \geqq m_n > 0$, $n > N$, *then* $\lim_{n \to \infty} \frac{1}{4}(b-a)^n m_n/n! = 0$.

(18) generally gives better results than those obtainable from (10).

Illustration. In the interval (0, 1)

$$h^{n+1} < \frac{E_n(e^{hx})}{E_n(e^x)} < h^{n+1}e^{-h} \qquad (h > 1),$$

$$\frac{k(k+1)\cdots(k+n)}{\alpha^{k-1}(n+1)!} < \frac{E_n((\alpha - x)^{-k})}{E_n((\alpha - x)^{-1})} < \frac{k(k+1)\cdots(k+n)}{(\alpha-1)^{k-1}(n+1)!}$$

$$(\alpha > 1;\, k > 1, \text{ arbitrary}).$$

With the notations of Theorem III and by the same reasoning, we now establish

THEOREM V. *Assume* (i) *neither* $\varphi^{(n)}(x)$, *nor* $f^{(n+1)}(x)$, *nor* $f^{(\nu)}(x)/E_n(f) \mp \varphi^{(\nu)}(x)/E_{n-1}(\varphi)$, $\nu > n$, *or* (ii) *neither* $\varphi^{(n+1)}(x)$, *nor* $f^{(n+1)}(x)$, *nor* $f^{(\nu)}(x)/E_n(f) - \varphi^{(\nu)}(x)/E_n(\varphi)$, $\nu > n + 1$, *changes sign in* (a, b). *Then correspondingly*: (i) *not more than* $n - \nu$ *points* $z_{j,n-1}$ *may fall between any two successive* $x_{i,n}$; (ii) *not more than* $n - \nu$ *points* $z_{i,n}$ *may fall between any two successive* $x_{i,n}$.

7. **Bounds for the highest coefficient of $\Pi_n(x; f)$; its sign.** Let

$$\Pi_n(x; f) \equiv \Pi_n(x) = \pi_{n,n}x^n + \pi_{n-1,n}x^{n-1} + \cdots. \tag{20}$$

Assume $E_{n-1}(f) > E_n(f)$. At the points of deviation $x_{j,n-1}$ $(j = 1, 2, \cdots, n+1)$, where $f(x) - \Pi_{n-1}(x)$ is of alternating signs, the difference

$$\Pi_n(x) - \Pi_{n-1}(x) \equiv f(x) - \Pi_{n-1}(x) - [f(x) - \Pi_n(x)] \tag{21}$$

has the same sign as $f(x) - \Pi_{n-1}(x)$; hence, $\Pi_n(x) - \Pi_{n-1}(x)$ vanishes between any two $x_{j,n-1}$, $x_{j+1,n-1}$,[6] so that $\Pi_{n-1}(x)$ *is in* (a, b) *a LIP, of degree not exceeding*

[6] Since $\Pi_n(x) - \Pi_{n-1}(x)$ has at most n real zeros, we thus learn once more that $E_{n-1}(f) \neq E_n(f)$ implies $f(x) - \Pi_{n-1}(x; f)$ has precisely $n + 1$ points of deviation, in accordance with (2) and (3).

$n - 1$, *for* $\Pi_n(x)$. By (1),

$$(22) \qquad \Pi_n(x) = \Pi_{n-1}(x) + \pi_{n,n} \prod_{j=1}^{n} (x - y_j) \qquad (x_{j,n-1} < y_j < x_{j+1,n-1}).$$

Moreover,

$$(23) \qquad | \Pi_n(x) - \Pi_{n-1}(x) | \leqq E_{n-1} + E_n, \quad a \leqq x \leqq b \qquad (E_n \equiv E_n(f)).$$

By the inequalities of Tchebycheff and W. Markoff [4], we get, taking hereafter, to simplify writing, $a = -1$, $b = 1$,

$$(24) \qquad \begin{aligned} | \pi_{n,n} | &\leqq 2^{n-1}[E_{n-1} + E_n], \\ | \pi_{n-1,n} - \pi_{n-1,n-1} | &\leqq 2^{n-2}[E_{n-1} + E_n], \cdots . \end{aligned}$$

This procedure, applicable to any continuous $f(x)$, is trivial. We can go further if we assume

$$(25) \qquad \textit{neither } f^{(n)}(x) \textit{ nor } f^{(n+1)}(x) \textit{ changes sign in } (-1, 1).$$

Then certainly $E_{n-1} \neq E_n$, hence, $E_{n-1} > E_n$, and we can use, in addition to (22), also (4), (6) and Corollary 1 to Theorem III. We readily show that *under* (25) $\pi_{n,n}$ *has the sign of* $f^{(n)}(x)$. In fact, assume $f^{(n)}(x) \geqq 0$ in $(-1, 1)$. Then, by (4), (6), (7) and (22),

$$\begin{aligned} \Pi_n(1) - \Pi_{n-1}(1) &= [f(1) - \Pi_{n-1}(1)] - [f(1) - \Pi_n(1)] \\ &= E_{n-1} \pm E_n = \pi_{n,n} \prod_{j=1}^{n} (1 - y_j) \quad (-1 < y_j < 1) \end{aligned}$$

($-$ or $+$, according as $f^{(n+1)}(x) \geqq 0$ or $\leqq 0$ in $(-1, 1)$), so that here $\pi_{n,n} > 0$. Similarly, if $f^{(n)}(x) \leqq 0$ in $(-1, 1)$, then, with the same y_j,

$$\Pi_n(1) - \Pi_{n-1}(1) = -E_{n-1} \mp E_n = \pi_{n,n} \prod_{j=1}^{n} (1 - y_j) < 0$$

($-$ or $+$, as above), so that $\pi_{n,n} < 0$.

Assume again, to be definite, $f^{(n)}(x) \geqq 0$ in $(-1, 1)$. We get from (22), by Corollary 1 to Theorem III, making use of the trigonometric polynomials $T_{n-1}(x)$, $T'_{n-1}(x)$ (see (8)), for $x = \pm 1$, $n > 1$,

$$(26) \qquad \begin{aligned} E_{n-1} + E_n &= \pi_{n,n} \prod_{j=1}^{n} (1 - y_j) < \frac{n-1}{2^{n-4}} \pi_{n,n}, \\ E_{n-1} - E_n &= \pi_{n,n} \prod_{j=1}^{n} (1 + y_j) < \frac{n-1}{2^{n-4}} \pi_{n,n}. \end{aligned}$$

Thus (see (24), (22))

$$(27) \qquad \frac{2^{n-4}}{n-1} (E_{n-1} + E_n) < \pi_{n,n} < 2^{n-1}(E_{n-1} + E_n),$$

$$(28) \qquad \frac{1}{2^{n-1}} < \prod_{j=1}^{n} (1 - y_j) = \frac{E_{n-1} + E_n}{\pi_{n,n}} < \frac{n-1}{2^{n-4}}.$$

It follows, by (23), that, *under condition* (25), $\pi_{n,n}^{-1}[\Pi_n(x) - \Pi_{n-1}(x)]$ *converges to zero uniformly in* $(-1, 1)$ *as* $n \to \infty$; more precisely,

$$\frac{1}{2^{n-1}} < \left| \frac{\Pi_n(x) - \Pi_{n-1}(x)}{\pi_{n,n}} \right| < \frac{n-1}{2^{n-4}} \qquad (-1 \leqq x \leqq 1). \tag{29}$$

As to other coefficients of $\Pi_n(x)$, we confine ourselves to the following remark. (22) gives

$$\frac{\pi_{n-1,n-1}}{\pi_{n,n}} - \frac{\pi_{n-1,n}}{\pi_{n,n}} = \sum_{j=1}^{n} y_j,$$

whence, again using Corollary 1 to Theorem III and $T'_{n-1}(x)$, we get

$$-2 < \frac{\pi_{n-1,n-1}}{\pi_{n,n}} - \frac{\pi_{n-1,n}}{\pi_{n,n}} < 2. \tag{30}$$

We close with the following remark. The foregoing is generally applicable to functions with *bounded derivatives* in the interval under consideration. If, for example, $f^{(n+1)}(x) \geqq m$ in (a, b), $m < 0$, we may introduce the function

$$\varphi(x) = f(x) + \frac{|m| + \alpha}{(n+1)!} x^{n+1} \qquad (\alpha \geqq 0).$$

Regarding the general case of $f(x)$ known only to be continuous in (a, b), observe ([3], pp. 91–92) that the approximate construction of $\Pi_n(x; f)$ can be reduced to that of $\Pi_n(x; P)$, where $P(x)$ is Weierstrass' approximation polynomial for $f(x)$ on (a, b), of sufficiently high degree. The foregoing discussion evidently applies to $P(x)$, also to $\varphi(x)$ introduced above, and we may then return to $f(x)$ using the property that both $\Pi_n(x; f)$ and $E_n(f)$, considered as operators on $f(x)$, *are continuous*.[7]

BIBLIOGRAPHY

1. S. BERNSTEIN, *Leçons sur les Propriétés Extrémales et la Meilleure Approximation des Fonctions Analytiques d'une Variable Réelle*, Paris, 1926.
2. S. BERNSTEIN, *On the Extremal Properties of Polynomials* (in Russian), Moscow, 1937.
3. E. BOREL, *Leçons sur les Fonctions des Variables Réelles et les Développements en Séries de Polynomes*, Paris, 1905.
4. W. MARKOFF, *On Functions Deviating the Least from Zero on a Given Interval* (in Russian), St. Pétersbourg, 1892; German translation, Mathematische Annalen, vol. 77 (1916), pp. 213–255.
5. W. STEKLOFF, *Quelques applications nouvelles de la théorie de fermeture au problème de représentation approchée des fonctions et au problème des moments*, Mémoires (Zapiski) de l'Académie des Sciences, St. Pétersbourg, Classe Phys.-Math., vol. 32(1914), pp. 1–74.
6. CH.-J. DE LA VALLÉE-POUSSIN, *Leçons sur l'Approximation des Fonctions d'une Variable Réelle*, Paris, 1919.

THE UNIVERSITY OF PENNSYLVANIA.

[7] After the present paper was completed, there came to the attention of the author a recent paper by E. J. Remes, *On some estimates of best approximation, and, in particular, on a fundamental theorem of de la Vallée-Poussin*, D. A. Grave Memorial Volume, pp. 235–244, Moscow, 1940, where, by a different method (divided differences), a formula is obtained equivalent to our formula (4), and—again in a different way—some of its most immediate consequences, like (10), are given.

THE DECOMPOSITION OF MEASURES

By Paul R. Halmos

1. **Introduction.** The main purpose of this paper is to prove a theorem on the decomposition of measures (Theorem 1 of §5)—a theorem which asserts that under certain hypotheses a measure space can be expressed as a direct sum of measure spaces. Although the result is of a certain independent interest, it is proposed chiefly as a new method: a tool to be used in the theory of measure-preserving transformations. In §6 we shall give, using this method, an easy proof of a theorem, due to von Neumann, on the decomposition of an arbitrary measure-preserving transformation into ergodic parts. In a subsequent paper we shall apply the method to the spectral theory of measure-preserving transformations to obtain a decomposition of an arbitrary measure-preserving transformation into parts which have either pure point spectrum or pure continuous spectrum.

2. **Measure spaces and separability.** Let Ω be any set of elements ω and let $\mathfrak{B}$ be a Borel field of subsets of Ω.[1] We suppose that on $\mathfrak{B}$ there is defined a measure m with $m(\Omega) = 1$: m is a non-negative, countably additive function of sets. We shall call Ω, together with $\mathfrak{B}$ and m, a *measure space*; when necessary we shall write $\Omega(\mathfrak{B}, m)$ for Ω, to emphasize the particular Borel field and measure under consideration. All the Borel fields we shall consider will be supposed to be (not necessarily proper) subfields of $\mathfrak{B}$, so that we may assume that the measure m is defined on them. If $\mathfrak{A}$ is a Borel field and A a set, $A \epsilon \mathfrak{A}$, we shall say that A is measurable ($\mathfrak{A}$); instead of measurable ($\mathfrak{B}$) we shall generally say measurable. A similar terminology will be used concerning the measurability of functions. We shall call the smallest Borel field containing a given collection of sets the Borel field spanned by them. For two sets, functions, transformations, etc., we shall use the symbol $\doteq$ to denote the fact that they are equal except possibly for a set of measure zero (i.e., equal *almost everywhere* or *a. e.*).[2] Two Borel fields $\mathfrak{A}_1$ and $\mathfrak{A}_2$ (both contained in $\mathfrak{B}$) will be called *equivalent*, in symbols $\mathfrak{A}_1 \cong \mathfrak{A}_2$, if to every set E in either one of them there corresponds a set F in the other so that $E \doteq F$.

There are two common notions of separability for measure spaces. $\Omega(\mathfrak{B}, m)$

Received February 12, 1941.

[1] For a definition of the notions *field, Borel field, strict separability*, etc., to be used throughout this paper, see [3], pp. 752–753. Numbers in brackets refer to the bibliography at the end of the paper.

[2] Thus if E and F are measurable sets we write $E \doteq F$ if $m(EF^{-1} + E^{-1}F) = 0$, where we use the notation E^{-1} for the complementary set $\Omega - E$.

is *strictly separable* if $\mathfrak{B}$ is spanned by a countable sequence of its sets. To define the other concept of separability we associate with every measure space $\Omega = \Omega(\mathfrak{B}, m)$ a complete metric space $\Omega^* = \Omega^*(\mathfrak{B}, m)$ as follows. The *points* of Ω^* are the *sets* $E \in \mathfrak{B}$; two sets E and F are the same point of Ω^* if $E \doteq F$; the distance between E and F is defined by $m(EF^{-1} + E^{-1}F)$. In terms of this metric space we define Ω to be *separable* if Ω^* is a separable metric space.[3]

LEMMA 1. *If $\Omega(\mathfrak{B}, m)$ is strictly separable, then it is separable; if $\Omega(\mathfrak{B}, m)$ is separable, then $\mathfrak{B}$ contains an equivalent subBorel field $\mathfrak{B}'$ such that $\Omega(\mathfrak{B}', m)$ is strictly separable.*[4]

Proof. If Ω is strictly separable and the sequence $B_1, B_2, \cdots (\in \mathfrak{B})$ spans $\mathfrak{B}$, then the field spanned by the B_n is countable and spans $\mathfrak{B}$. A well-known approximation theorem shows that Ω is then separable.[5]

Suppose on the other hand that Ω is separable. Let $B_1, B_2, \cdots$ be a countable everywhere dense sequence in Ω^*, and let $\mathfrak{B}'$ be the Borel field spanned by the B_n. Then the space $\Omega^*(\mathfrak{B}', m)$ is a subset of $\Omega^*(\mathfrak{B}, m)$ and since it is a complete metric space, it is a closed subset of $\Omega^*(\mathfrak{B}, m)$. Since, finally, $\Omega^*(\mathfrak{B}', m)$ is everywhere dense in $\Omega^*(\mathfrak{B}, m)$, it follows that $\Omega^*(\mathfrak{B}', m) = \Omega^*(\mathfrak{B}, m)$. But the last statement is equivalent to $\mathfrak{B}' \cong \mathfrak{B}$.

In applications it is usually permissible to assume separability; we proved Lemma 1 in order to justify our assumption, to be made later, of strict separability. Using Lemma 1, we can easily prove, by means of the standard methods of measure theory, that every function or transformation which is measurable ($\mathfrak{B}$) is equal almost everywhere to one measurable ($\mathfrak{B}'$).

3. Direct sums of measure spaces. Let $X = X(\mathfrak{X}, \mu)$ be a measure space and suppose that to each $x \in X$ there corresponds a measure space $Y_x = Y_x(\mathfrak{Y}_x, \nu_x)$. Denote by Ω the set of all pairs (x, y) with $x \in X$ and $y \in Y_x$. In order to avoid the introduction of any more new notation than necessary we shall write Y_{x_0} for the set of all pairs (x_0, y) (points of Ω) for which $y \in Y_{x_0}$. We shall say that a set $E \subseteq \Omega$ depends on x alone (or E is an x-set) if for every $x \in X$ either $E \cdot Y_x$ is empty or $E \cdot Y_x = Y_x$. Once more we carry over to Ω the notation of the component spaces: we denote by $\mathfrak{X}$ the class of all x-sets E for which the set of points x such that $E \cdot Y_x = Y_x$ is a measurable subset of X.

Let $\mathfrak{S}$ be the class of all sets $E \subseteq \Omega$ for which $E \cdot Y_x$ is a measurable subset of Y_x for each x, and for which $\nu_x(E \cdot Y_x)$ is a measurable function of x. ($\mathfrak{S}$ is not in general a Borel field: it is a normal class. See [7], p. 83.) It is clear

[3] It is easy to see that Ω is separable in this sense if and only if all the function spaces $L_p(\Omega)$, $p \geqq 1$, are separable.

[4] The unit interval, with the Lebesgue measurable sets and Lebesgue measure, is separable but not strictly separable. Concerning the hereditary properties of separability the following comments are relevant. If $\Omega(\mathfrak{B}, m)$ is separable and $\mathfrak{B}'$ is a Borel field, $\mathfrak{B}' \subseteq \mathfrak{B}$, then $\Omega(\mathfrak{B}', m)$ is separable. The analogue of this result for strict separability is not true. Thus the unit interval, with the Borel sets and Lebesgue measure, is strictly separable, but the Borel field spanned by the sets consisting of exactly one point is not.

[5] See, for example, [4], p. 4.

that $\mathfrak{X} \subseteq \mathfrak{S}$. For every $E \in \mathfrak{S}$ we define

$$(*)\qquad m(E) = \int_X \nu_x(E \cdot Y_x)\, d\mu(x):$$

this integral exists since $\nu_x(E \cdot Y_x)$ is by hypothesis measurable, and by the definition of a measure space, $0 \leqq \nu_x(E \cdot Y_x) \leqq 1$. If $\mathfrak{B}$ is a Borel field of subsets of Ω such that $\mathfrak{X} \subseteq \mathfrak{B} \subseteq \mathfrak{S}$ and such that the set function m is a measure on $\mathfrak{B}$, then we say that the measure space $\Omega(\mathfrak{B}, m)$ is a *direct sum* of the spaces Y_x formed with respect to X. We note that $\Omega(\mathfrak{B}, m)$ is not uniquely determined by the spaces X and Y_x: it depends also on the choice of the Borel field $\mathfrak{B}$. The definition is not, however, vacuous: the Borel field $\mathfrak{X}$ can always be chosen for $\mathfrak{B}$. For we have already observed that $\mathfrak{X} \subseteq \mathfrak{S}$, and since $(*)$ guarantees that for $E \in \mathfrak{X}$, $m(E) = \mu(E)$, we know also that m is a measure on $\mathfrak{X}$. Without any further hypotheses it does not seem possible to prove that a $\mathfrak{B}$ different from $\mathfrak{X}$ always exists: in the applications, of course, we shall generally start with a Borel field $\mathfrak{B}$ much larger than $\mathfrak{X}$.

4. **The relative density function.** Let $\Omega = \Omega(\mathfrak{B}, m)$ be a measure space, and let $\mathfrak{A}$ be a Borel field, $\mathfrak{A} \subseteq \mathfrak{B}$. For an arbitrary fixed set $B \in \mathfrak{B}$ the set function $n(A) = m(AB)$ is a finite measure which vanishes whenever $m(A)$ vanishes. Hence we may apply the Radon-Nikodym theorem[6] to this set function considered on the measure space $\Omega(\mathfrak{A}, m)$ (*not* $\Omega(\mathfrak{B}, m)$), and find a *relative density function* $\delta(B, \omega)$ such that

(i) $\delta(B, \omega)$ *is measurable* $(\mathfrak{A})$, *as a function of* $\omega \in \Omega$ *for every* B; and

(ii) $m(AB) = \int_A \delta(B, \omega)\, dm(\omega)$ *for every* $A \in \mathfrak{A}$ *and* $B \in \mathfrak{B}$.[7]

$\delta(B, \omega)$ is uniquely determined by the conditions (i) and (ii) in the sense that if $\delta'(B, \omega)$ also satisfies these conditions, then, for every B, $\delta(B, \omega) \doteq \delta'(B, \omega)$. The set of measure zero on which δ and δ' differ belongs to $\mathfrak{A}$ and may, of course, depend on B. Before discussing the dependence of $\delta(B, \omega)$ on B for fixed ω, we wish to emphasize the following fact. The condition (ii) alone does not, in general, serve to determine $\delta(B, \omega)$. (ii) is satisfied, for example, by the characteristic function of B, but unless $B \in \mathfrak{A}$ this function does not satisfy (i). (The relation of $\delta(B, \omega)$ to the characteristic function of B will be discussed in more detail below.) As a very illuminating example, illustrating this point as well as all others to follow, the reader is advised to keep in mind the unit square in the rôle of Ω (with the class of all Borel sets and Lebesgue measure for $\mathfrak{B}$ and m respectively) and take for $\mathfrak{A}$ the class of all Borel sets depending on x alone. If $\varphi_B(x, y)$ is the characteristic function of B, then it is not $\varphi_B(x, y)$ but rather the function $\int_0^1 \varphi_B(x, y)\, dy$ that plays the rôle of $\delta(B, \omega)$.

[6] See [7], p. 36.

[7] The idea of introducing the relative density function is a familiar one in probability theory, where functions similar to our density are known as conditional probabilities. See, for example, [5], pp. 41–44.

We now mention two further properties of $\delta(B, \omega)$ concerning the way in which it depends on B for fixed ω.

(iii) *If $\Omega(\mathfrak{B}, m)$ is strictly separable, there exists a function $\delta'(B, \omega)$ with properties* (i) *and* (ii) *of $\delta(B, \omega)$ (so that $\delta'(B, \omega) \doteq \delta(B, \omega)$), and a set $A' \in \mathfrak{A}$ with $m(A') = 0$, such that for ω fixed and not in A', $\delta'(B, \omega)$ is a measure on $\mathfrak{B}$.*

This result is due to Doob (see [2], pp. 95–98). The statement and proof of his theorem are given in a notation very different from ours, but the proof applies to our case. In the statement and proof of the following property we retain the notation established in (iii) and use the result quoted there. For any set E we write $\varphi(E, \omega)$ for the characteristic function of E.

(iv) *If $\Omega(\mathfrak{B}, m)$ and $\Omega(\mathfrak{A}, m)$ are both strictly separable, then there exists a set $A'' \in \mathfrak{A}$, with $m(A'') = 0$, such that for ω not in A'' and A arbitrary in $\mathfrak{A}$, $\delta(A, \omega) = \varphi(A, \omega)$.*

Proof. Let $A_1, A_2, \cdots$ be a countable sequence of sets of $\mathfrak{A}$ which span $\mathfrak{A}$. There is no loss of generality in assuming that the sets A_n form a *field*, i.e., that they are a system closed under the processes of forming finite sums, products, and complements. Since for each $n = 1, 2, \cdots, \varphi(A_n, \omega)$ is measurable ($\mathfrak{A}$) and for every $A \in \mathfrak{A}$, $m(AA_n) = \int_A \varphi(A_n, \omega)\, dm(\omega)$—in other words, since $\varphi(A_n, \omega)$ has the two defining properties (i) and (ii) of $\delta(A_n, \omega)$—there must exist a set $A'_n \in \mathfrak{A}$, with $m(A'_n) = 0$, such that for $\omega \in' A'_n$, $\delta'(A_n, \omega) = \varphi(A_n, \omega)$. We write

$$A'' = A' + A'_1 + A'_2 + \cdots$$

(where A' is the set described in (iii)). If ω is not in A'', then, in the first place, $\delta'(B, \omega)$ is a measure on $\mathfrak{B}$, and therefore on $\mathfrak{A}$, and in the second place (since $\delta'(A_n, \omega) = \varphi(A_n, \omega)$) its value for any set A_n is one or zero according as $\omega \in A_n$ or $\omega \in' A_n$. Since the A_n form a field which spans $\mathfrak{A}$, the measure $\delta'(A_n, \omega)$, considered, for the moment, as if it were defined for only the A_n, has one and only one extension to $\mathfrak{A}$.[8] The fact that $\delta'(A, \omega)$ and $\varphi(A, \omega)$ are both measures on $\mathfrak{A}$, which agree if $A = A_n$, implies therefore that $\delta'(A, \omega) = \varphi(A, \omega)$ for all $A \in \mathfrak{A}$. Since for $\omega \in' A''$, $\delta'(B, \omega) = \delta(B, \omega)$ for all $B \in \mathfrak{B}$, the assertion (iv) is proved.

5. **The decomposition theorem.** In this section we shall assume that $\Omega(\mathfrak{B}, m)$ is a strictly separable measure space and that $\mathfrak{A}$ is a Borel field, $\mathfrak{A} \subsetneq \mathfrak{B}$, such that $\Omega(\mathfrak{A}, m)$ is also strictly separable. In terms of $\Omega = \Omega(\mathfrak{B}, m)$ and $\mathfrak{A}$ we shall now describe certain other measure spaces which will serve to express Ω as a direct sum.

We observe first that the Borel field $\mathfrak{A}$, being strictly separable, is atomic: i.e., there exists in $\mathfrak{A}$ a system $\mathfrak{A}_0$ of sets such that if $A_0 \in \mathfrak{A}_0$ then no proper subset of A_0 is in $\mathfrak{A}$ and such that every $A \in \mathfrak{A}$ is a sum (not necessarily countable) of sets in $\mathfrak{A}_0$. For let $A_1, A_2, \cdots$ be a countable sequence of sets of $\mathfrak{A}$ which

[8] See [5], pp. 15–16.

span $\mathfrak{A}$ and let $\mathfrak{A}_0$ be the class of all non-empty sets of the form $\prod_{i=1}^{\infty} A_i^{\epsilon_i}$, where $\epsilon_i = \pm 1$, A_i^{+1} denotes A_i, and A_i^{-1} denotes, as usual, the complement of A_i. It is clear, moreover, that if $f(\omega)$ is any function measurable ($\mathfrak{A}$), then on any atom A_0 of $\mathfrak{A}$ (i.e., for any $A_0 \epsilon \mathfrak{A}_0$) $f(\omega)$ is constant.

Changing to a more suggestive notation, we denote by X the set of all those atoms of $\mathfrak{A}$ which are disjoint from the exceptional set A'' described in (iv) of §4, and we proceed to define a concept of measurability and measure in X. A subset of X will be called measurable if the sum of the atoms in it is a set measurable ($\mathfrak{A}$), and we assign as measure μ to this set the measure m of the corresponding set of $\mathfrak{A}$. For every $x \epsilon X$ we denote by Y_x the atom corresponding to x: Y_x is a subset, belonging to $\mathfrak{A}_0$, of Ω, whereas x is to be thought of as an abstract entity, an element of a measure space X. The concepts of measurability and measure are defined in Y_x as follows. A subset of Y_x is measurable if it is of the form $Y_x \cdot B$ with $B \epsilon \mathfrak{B}$; we choose an arbitrary $\omega \epsilon Y_x$ and define $\nu_x(Y_x \cdot B) = \delta(B, \omega)$. Concerning this definition we make the following two comments.

I. Since $\delta(B, \omega)$ is measurable ($\mathfrak{A}$) its value on any atom of $\mathfrak{A}$ is a constant independent of ω. Hence the definition of $\nu_x(Y_x \cdot B)$ does not depend on the particular $\omega \epsilon Y_x$ that was chosen.

II. It may happen that for two different sets B' and B'' ($\epsilon \mathfrak{B}$) $Y_x \cdot B' = Y_x \cdot B''$: we have to show that in this case $\omega \epsilon Y_x$ implies $\delta(B', \omega) = \delta(B'', \omega)$. Since we have already excluded the exceptional set A'' of (iv), §4, we know that (with ω fixed in Y_x) $\delta(B, \omega)$ is a measure on $\mathfrak{B}$, whose value for a set $A \epsilon \mathfrak{A}$ is one or zero according as $\omega \epsilon A$ or $\omega \epsilon' A$. In particular therefore $\delta(Y_x, \omega) = 1$. According to a result of Doob[9] a measure (such as $\delta(B, \omega)$) uniquely determines a measure on a measurable subspace (such as Y_x) in the way in which we described, if and only if the measure of the subspace is one. Since we have shown that in our case this condition is satisfied, we may conclude that our definition of $\nu_x(Y_x \cdot B)$ depends only on the intersection $Y_x \cdot B$ and not on the particular set B.

If we compare the defining property (ii) of $\delta(B, \omega)$ with the relation (*) used in the description of direct sums, it is clear that we may sum up our results in the following decomposition theorem.

THEOREM 1. *If $\Omega = \Omega(\mathfrak{B}, m)$ is a strictly separable measure space and $\mathfrak{A}$ a Borel field, $\mathfrak{A} \subseteq \mathfrak{B}$, such that $\Omega(\mathfrak{A}, m)$ is also strictly separable then, except possibly for a set $A'' \epsilon \mathfrak{A}$ of measure zero, Ω is a direct sum of measure spaces Y_x formed with respect to a measure space X in such a way that the Borel field $\mathfrak{X}$ of all measurable x-sets coincides with the given Borel field $\mathfrak{A}$.*

6. **Measure-preserving transformations.** We shall call a one-to-one transformation of a measure space $\Omega(\mathfrak{B}, m)$ on itself *measure preserving* if the image

[9] [1], pp. 109–110. Doob's result is more general than what we need here: it requires only that the *outer* measure of the subspace be one.

of a measurable set under both T and T^{-1} is a measurable set of the same measure. A set E is an *invariant set* (of T) if $TE = E$.[10] The transformation T is *ergodic* (or *metrically transitive*) if for every measurable invariant set E we have either $m(E) = 0$ or $m(E) = 1$. An equivalent, and for our purposes more convenient, form of this definition is the following. T is ergodic if for every measurable set E the measure of the set E^*, defined by $E^* = \sum_{n=-\infty}^{\infty} T^n E$, is one or zero.

Before stating von Neumann's decomposition theorem, we make a remark on separability. If T is a measure-preserving transformation on $\Omega(\mathcal{B}, m)$, the collection $\mathcal{A}$ of invariant sets is easily verified to be a Borel field. Even if we knew that $\Omega(\mathcal{B}, m)$ (and therefore $\Omega(\mathcal{A}, m)$) is separable, we would not be able to apply Theorem 1, for there we assumed strict separability. We may, however, apply Lemma 1 and find a countable sequence of sets $C_1, C_2, \cdots$ which span a Borel field $\mathcal{B}'$ equivalent to $\mathcal{B}$ and which are such that at the same time those C_n which belong to $\mathcal{A}$ span a Borel field $\mathcal{A}'$ equivalent to $\mathcal{A}$. Then the countable collection of sets of the form $T^i C_j$ $(i, j = 1, 2, \cdots)$ span a Borel field $\mathcal{B}''$ also equivalent to $\mathcal{B}$, and we may apply Theorem 1 to $\mathcal{B}''$ and $\mathcal{A}'$. In order to avoid a superfluity of notation we shall not, therefore, assume separability, in the statement of Theorem 2 below, but we shall assume that both Borel fields we shall consider are already strictly separable.

THEOREM 2. *If $\Omega = \Omega(\mathcal{B}, m)$ is a strictly separable measure space and if T is a measure-preserving transformation on Ω whose invariant sets form a Borel field $\mathcal{A}$ such that $\Omega(\mathcal{A}, m)$ is strictly separable, then, except possibly for a set $A'' \in \mathcal{A}$ of measure zero, Ω is a direct sum of measure spaces Y_x formed with respect to a measure space X, in such a way that each Y_x is an invariant set of T and, except possibly for an x-set of measure zero, the transformation T_x induced by T on Y_x is an ergodic measure-preserving transformation.*[11]

Proof. An application of Theorem 1 shows immediately that Ω is a direct sum and that each Y_x is invariant. It follows that T induces a one-to-one transformation T_x of each Y_x on itself: the only thing that has to be proved is that T_x is measure preserving and ergodic (for almost all x).

Since for any measurable set B and any x, $T(Y_x \cdot B) = Y_x \cdot TB$, and $T^{-1}(Y_x \cdot B) = Y_x \cdot T^{-1}B$, it is clear that for any measurable subset E of Y_x both $T_x \cdot E$ and $T_x^{-1}E$ are measurable.

[10] An invariant set is sometimes defined as a set E for which $TE \doteq E$. The present definition is a little more convenient for our purposes.

[11] This theorem was proved by von Neumann ([6], pp. 601–618) not for a single transformation but for a one-parameter group of transformations (a flow), and not on an arbitrary measure space but on a space satisfying certain topological restrictions (an m-space), which in his case need not have finite measure. Our proof applies, with the usual changes necessitated by measurability difficulties, to the case of a flow on a space of not necessarily finite measure, and thus our result is a generalization of von Neumann's theorem inasmuch as it does not use topological considerations.

Let $B_1, B_2, \cdots$ be a countable sequence of measurable sets which spans $\mathfrak{B}$; as in the proof of (iv) in §4 we assume that the B_n form a field. For any $n = 1, 2, \cdots$ and any $A \in \mathfrak{A}$ we have

$$\int_A \delta(TB_n, \omega)\, dm(\omega) = m(A \cdot TB_n) = m(TA \cdot TB_n)$$

$$= m(A \cdot B_n) = \int_A \delta(B_n, \omega)\, dm(\omega).$$

It follows that except for ω in a set $\bar{A}_n \in \mathfrak{A}$ of measure zero $\delta(TB_n, \omega) = \delta(B_n, \omega)$. Hence if ω does not lie in the set $A'' + \bar{A}_1 + \bar{A}_2 + \cdots$ (where A'' is the set described in (iv), §4), $\delta(TB, \omega)$ and $\delta(B, \omega)$ are both measures on $\mathfrak{B}$ which agree on a field which spans $\mathfrak{B}$: it follows that they agree for all $B \in \mathfrak{B}$.[12] Thus we see, remembering the definition of measure in Y_x, that for almost all x, T_x is measure preserving.

Using the results of the preceding paragraph and of §4, we know that if ω_0 is taken outside a certain fixed set of measure zero, then $\delta(B, \omega_0)$ is a measure on $\mathfrak{B}$, $\delta(A, \omega_0) = \varphi(A, \omega_0)$ for $A \in \mathfrak{A}$, and $\delta(TB, \omega_0) = \delta(B, \omega_0)$. In other words $\Omega(\mathfrak{B}, \delta(B, \omega_0))$ is a measure space, T is a measure-preserving transformation on this space, and the fact that for every $B \in \mathfrak{B}$, $B^* = \sum_{n=-\infty}^{\infty} T^n B \in \mathfrak{A}$ (i.e., B^* is an invariant set of T), implies that $\delta(B^*, \omega_0) = \varphi(B^*, \omega_0) = 1$ or 0. Recalling once more the definition of measure in Y_x and the definition of ergodicity, we infer that T_x is ergodic. This completes the proof of Theorem 2.

Bibliography

1. J. L. Doob, *Stochastic processes depending on a continuous parameter*, Transactions of the American Mathematical Society, vol. 42(1937), pp. 107–140.
2. J. L. Doob, *Stochastic processes with an integral-valued parameter*, Transactions of the American Mathematical Society, vol. 44(1938), pp. 87–150.
3. J. L. Doob, *One-parameter families of transformations*, this Journal, vol. 4(1938), pp. 752–774.
4. E. Hopf, *Ergodentheorie*, Berlin, 1937.
5. A. Kolmogoroff, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlin, 1933.
6. J. v. Neumann, *Zur Operatorenmethode in der klassischen Mechanik*, Annals of Mathematics, vol. 33(1932), pp. 587–642.
7. S. Saks, *Theory of the Integral*, Warszawa, 1937.

The Institute for Advanced Study.

[12] The proof runs as follows: the collection of sets on which two measures are equal forms a normal class (see [7], p. 83). If a normal class contains a field, it contains the Borel field spanned by it.

FUNCTIONS HAVING SUBHARMONIC LOGARITHMS

By E. F. Beckenbach

1. **Introduction.** A function $x(u, v)$, defined in a domain D (non-null connected open set), such that $x(u, v) < +\infty$, is said to be *subharmonic*[1] in D provided it satisfies the following conditions:

1.1. $x(u, v)$ is upper semi-continuous in D;

1.2. for every domain D' comprised together with its boundary B' in D, and for every function $h(u, v)$, harmonic in D', continuous in $D' + B'$, and satisfying $h(u, v) \geqq x(u, v)$ on B', we have also $h(u, v) \geqq x(u, v)$ in D'.

A function $x(u, v)$, which is upper semi-continuous and $\not\equiv -\infty$ in D, is subharmonic there if and only if it satisfies

$$x(u_0, v_0) \leqq \frac{1}{2\pi} \int_0^{2\pi} x(u_0 + \rho \cos \varphi, v_0 + \rho \sin \varphi)\, d\varphi$$

for every point (u_0, v_0) in D and for every ρ such that the circular disc $(u - u_0)^2 + (v - v_0)^2 \leqq \rho^2$ is comprised in D. A function $x(u, v)$, having continuous second derivatives in D, is subharmonic there if and only if its Laplacian satisfies the inequality

$$\Delta x(u, v) \equiv \frac{\partial^2 x}{\partial u^2} + \frac{\partial^2 x}{\partial v^2} \geqq 0.$$

A function $p(u, v)$, defined in a domain D, is said to be[2] *of class PL* in D provided $p(u, v) \geqq 0$ and $\log p(u, v)$ is subharmonic there.

A function of class PL necessarily is subharmonic, but the converse is by no means true. Indeed, a non-negative function $p(u, v)$ is of class PL if and only if $[p(u, v)]^\alpha$ is subharmonic for all choices of the positive constant α. Other criteria for functions of class PL are contained in the following Theorems A and B.

Received February 19, 1941.

[1] For a discussion of subharmonic functions, see T. Radó, *Subharmonic Functions*, Berlin, 1937. There, as usual, subharmonic functions are defined to be functions which satisfy $x(u, v) \not\equiv -\infty$, in addition to the conditions we have here given; but, for example, $\log |f(u + iv)|$, where $f(z)$ is analytic, is subharmonic, and we prefer not to exclude the important special case $f(z) \equiv 0$. The discussions in the present paper would not, however, be affected by the restriction $x(u, v) \not\equiv -\infty$.

[2] See E. F. Beckenbach and T. Radó, *Subharmonic functions and minimal surfaces* and *Subharmonic functions and surfaces of negative curvature*, Transactions of the American Mathematical Society, vol. 35(1933), pp. 648–661 and 662–674 for definition and applications of functions of class PL.

THEOREM A. *If $p(u, v) \geqq 0$ in D, then $p(u, v)$ is of class PL in D if and only if $e^{\alpha u+\beta v}p(u, v)$ is subharmonic there for every choice of the real constants α, β.*

Theorem A was proved by Montel on the assumption that $p(u, v)$ has continuous second derivatives, and by Radó in the general case. Further results suggested by Theorem A have been obtained by Saks and Beckenbach.[3]

In the present note, we shall give yet another extension (see §2) of Theorem A; namely, we shall introduce the parameters α, β in quite a different way.

THEOREM B. *If $p(u, v)$ is continuous and $\geqq 0$ in D, then $p(u, v)$ is of class PL in D if and only if*

$$\frac{1}{\pi\rho^2}\iint_{\xi^2+\eta^2<\rho^2}[p(u_0+\xi, v_0+\eta)]^2\,d\xi\,d\eta \leqq \left[\frac{1}{2\pi}\int_0^{2\pi} p(u_0+\rho\cos\varphi, v_0+\rho\sin\varphi)\,d\varphi\right]^2$$

holds for every point (u_0, v_0) in D and for every ρ such that the circular disc $(u-u_0)^2+(v-v_0)^2 \leqq \rho^2$ is comprised in D.

The inequality in Theorem B is essentially the isoperimetric inequality, $a \leqq l^2/(4\pi)$; it follows from this theorem that the isoperimetric inequality characterizes surfaces of negative curvature.[4]

While Theorem B gives a characterization of a function of class PL in terms of itself, in what follows we shall present (see §3) an analogous theorem giving a characterization of these same functions in terms of other functions of class PL.

2. A two-parameter characterization of functions of class PL. We observe that in Theorem A, the function

$$e^{\alpha u+\beta v} \equiv \left| e^{(\alpha-i\beta)(u+iv)} \right|$$

is just barely of class PL, for its logarithm, $\alpha u+\beta v$, is harmonic. But a similar remark might be applied to the absolute value of any other analytic function $f(z)$; and the fact that $\alpha u+\beta v$ is actually linear in u, v appears to be unimportant in this connection.

THEOREM 1. *If $p(u, v) \geqq 0$ in D, then $p(u, v)$ is of class PL in D if and only if $[(u-\alpha)^2+(v-\beta)^2]p(u, v)$ is subharmonic there for every choice of the real constants α, β.*

Necessity. Let $p(u, v)$ be of class PL. We observe that also

$$[(u-\alpha)^2+(v-\beta)^2] \equiv |(u+iv)-(\alpha+i\beta)|^2$$

is of this class. It follows from the definition of the class PL and the fact that the sum of two subharmonic functions is again a subharmonic function, that

[3] P. Montel, *Sur les fonctions convexes et les fonctions sousharmoniques*, Journal de Mathématiques, (9), vol. 7(1928), pp. 29–60; T. Radó, *Remarque sur les fonctions subharmoniques*, Comptes Rendus, Paris, vol. 186(1928), pp. 346–348; S. Saks, *On subharmonic functions*, Acta Szeged, vol. 5(1930–32), pp. 187–193; E. F. Beckenbach, *On subharmonic functions*, this Journal, vol. 1(1935), pp. 480–483.

[4] E. F. Beckenbach and T. Radó, loc. cit., pp. 662–674.

$[(u - \alpha)^2 + (v - \beta)^2]p(u, v)$ is of class PL. Indeed, the product of any two functions of class PL obviously is again of class PL. Then

$$[(u - \alpha)^2 + (v - \beta)^2]p(u, v)$$

is subharmonic, since a function of class PL necessarily is subharmonic.

Sufficiency. For the non-negative function $p(u, v)$, let

$$[(u - \alpha)^2 + (v - \beta)^2]p(u, v)$$

be subharmonic for all choices of the real constants α, β. We assume first that $p(u, v)$ is positive and has continuous second derivatives in D. A computation yields

$$\Delta\{[(u - \alpha)^2 + (v - \beta)^2]p(u, v)\}$$
$$\equiv [(u - \alpha)^2 + (v - \beta)^2]\Delta p + 4p_u(u - \alpha) + 4p_v(v - \beta) + 4p,$$

whence, by assumption, the quadratic in $(u - \alpha)$, $(v - \beta)$ satisfies

$$[(u - \alpha)^2 + (v - \beta)^2]\Delta p + 4p_u(u - \alpha) + 4p_v(v - \beta) + 4p \geqq 0$$

for all points (u, v) in D and for every choice of the real constants α, β. It follows immediately that the discriminant satisfies

$$p\Delta p - (p_u^2 + p_v^2) \geqq 0.$$

But

$$p\Delta p - (p_u^2 + p_v^2) \equiv p^2\Delta \log p,$$

so that $p(u, v)$ is of class PL.

We assume now that $p(u, v)$ is only non-negative and has continuous second derivatives in D. Then the function $p(u, v) + \epsilon$ is positive and has continuous second derivatives in D for every $\epsilon > 0$. Further, since

$$[(u - \alpha)^2 + (v - \beta)^2]p(u, v)$$

is subharmonic, it follows that also

$$[(u - \alpha)^2 + (v - \beta)^2][p(u, v) + \epsilon]$$

is subharmonic. Hence, according to the preceding paragraph, $p(u, v) + \epsilon$ is of class PL, and therefore its uniform limit as $\epsilon \to 0$, $p(u, v)$, is of class PL.

Suppose finally that $p(u, v)$ is given to satisfy only the conditions specified in the theorem. We shall show first that $p(u, v)$ is subharmonic. That $p(u, v)$ is upper semi-continuous follows from the fact that $[(u - \alpha)^2 + (v - \beta)^2]$ is continuous and $[(u - \alpha)^2 + (v - \beta)^2]p(u, v)$, being subharmonic, is upper semi-continuous. Further, from the subharmonic character of

$$[(u - \alpha)^2 + (v - \beta)^2]p(u, v),$$

it follows[5] that

$$
\begin{aligned}
(1)\qquad [(u_0 - \alpha)^2 + (v_0 - \beta)^2]p(u_0, v_0) \leqq \frac{1}{2\pi}\int_0^{2\pi} [(u_0 - \alpha + \rho\cos\varphi)^2 \\
+ (v_0 - \beta + \rho\sin\varphi)^2]p(u_0 + \rho\cos\varphi, v_0 + \rho\sin\varphi)\, d\varphi
\end{aligned}
$$

holds for all circular discs, with center (u_0, v_0) and radius ρ in D, and for every choice of the real constants α, β. For a fixed $(u_0, v_0; \rho)$, let

$$(u_0 - \alpha)^2 + (v_0 - \beta)^2 = M^2, \qquad M > 0.$$

It follows from (1) that

$$p(u_0, v_0) \leqq \left(\frac{M+\rho}{M}\right)^2 \left(\frac{1}{2\pi}\right) \int_0^{2\pi} p(u_0 + \rho\cos\varphi, v_0 + \rho\sin\varphi)\, d\varphi.$$

Since M is arbitrarily large, we obtain

$$p(u_0, v_0) \leqq \frac{1}{2\pi}\int_0^{2\pi} p(u_0 + \rho\cos\varphi, v_0 + \rho\sin\varphi)\, d\varphi,$$

whence it follows[6] that $p(u, v)$ is subharmonic.

Since $p(u, v)$ is subharmonic and $\not\equiv -\infty$, it is summable,[7] so that we may consider the averaging functions,

$$p_r^{(0)}(u, v) \equiv p(u, v),$$

$$p_r^{(k)}(u, v) = \frac{1}{\pi r^2}\iint_{\xi^2+\eta^2<r^2} p_r^{(k-1)}(u + \xi, v + \eta)\, d\xi\, d\eta \qquad (k = 1, 2, \cdots).^8$$

From the fact that $p(u, v)$ satisfies (1), it follows that the averaging functions $p_r^{(k)}(u, v)$ also satisfy (1). Thus

$$
\begin{aligned}
[(u_0 - \alpha)^2 + (v_0 - \beta)^2]p_r^{(1)}(u_0, v_0) = \frac{1}{\pi r^2}\iint_{\xi^2+\eta^2<r^2} \{[(u_0 + \xi) - (\alpha + \xi)]^2 \\
+ [(v_0 + \eta) - (\beta + \eta)]^2\} p(u_0 + \xi, v_0 + \eta)\, d\xi\, d\eta \\
\leqq \frac{1}{\pi r^2}\iint_{\xi^2+\eta^2<r^2} \left\{\frac{1}{2\pi}\int_0^{2\pi} [(u_0 - \alpha + \rho\cos\varphi)^2 + (v_0 - \beta + \rho\sin\varphi)^2] \right. \\
\left. \cdot p(u_0 + \xi + \rho\cos\varphi, v_0 + \eta + \rho\sin\varphi)\, d\varphi \right\} d\xi\, d\eta
\end{aligned}
$$

[5] F. Riesz, *Sur les fonctions subharmoniques et leur rapport à la théorie du potentiel*, I, Acta Mathematica, vol. 48(1926), pp. 329–343.

[6] J. E. Littlewood, *On the definition of a subharmonic function*, Journal of the London Mathematical Society, vol. 2(1927), pp. 189–192.

[7] For justification of various statements made in this paragraph, see T. Radó, *Subharmonic Functions*, §§1.10, 2.21 and 3.8.

[8] Of course $p_r^{(k)}(u, v)$ can be defined thus for only a subdomain $D_r^{(k)}$ of D, but this is of no consequence since r is arbitrarily small.

$$= \frac{1}{2\pi}\int_0^{2\pi}\left\{\frac{1}{\pi r^2}\iint_{\xi^2+\eta^2<r^2}[(u_0-\alpha+\rho\cos\varphi)^2+(v_0-\beta+\rho\sin\varphi)^2]\right.$$
$$\left.\cdot p(u_0+\xi+\rho\cos\varphi, v_0+\eta+\rho\sin\varphi)\,d\xi\,d\eta\right\}d\varphi$$
$$= \frac{1}{2\pi}\int_0^{2\pi}[(u_0-\alpha+\rho\cos\varphi)^2+(v_0-\beta+\rho\sin\varphi)^2]$$
$$\cdot p_r^{(1)}(u_0+\rho\cos\varphi, v_0+\rho\sin\varphi)\,d\varphi,$$

so that $p_r^{(1)}(u, v)$ satisfies (1). Similarly $p_r^{(2)}(u, v)$ and $p_r^{(3)}(u, v)$ satisfy (1). Further, $p_r^{(3)}(u, v) \geqq 0$ and has continuous second derivatives. Therefore, as above, $p_r^{(3)}(u, v)$ is of class PL. But we have shown that $p(u, v)$ is subharmonic, so that $p_r^{(3)}(u, v) \searrow p(u, v)$ as $r \to 0$. Since the limit of a decreasing sequence of functions of class PL is again of class PL, it follows that $p(u, v)$ is of class PL.

The class of functions of class PL is invariant under conformal mappings of the (u, v)-domain of definition; accordingly, we have the following corollary.

COROLLARY. *Let $U + iV = f(u + iv)$ be a given analytic function, which is not identically constant, in D. If $p(u, v) \geqq 0$ in D, then $p(u, v)$ is of class PL in D if and only if*

$$\{[U(u, v)-\alpha]^2+[V(u, v)-\beta]^2\}p(u, v) \equiv |f(u+iv)-(\alpha+i\beta)|^2 p(u, v)$$

is subharmonic there for every choice of the real constants α, β.

We note that Theorem 1 still holds if we replace

$$(u-\alpha)^2+(v-\beta)^2$$

by the distance function

$$[(u-\alpha)^2+(v-\beta)^2]^{\frac{1}{2}}.$$

It follows that a similar remark may be made concerning the corollary.

3. An isoperimetric characterization of functions of class PL. We have stated the following Theorem 2 in terms of *all* continuous functions $q(u, v)$ of class PL in order to have the necessity of condition (2) general; but for the sufficiency we note that we need assume only that (2) holds for a suitable two-parameter subclass of functions of class PL, for example, the subclass $e^{\alpha u+\beta v}$ of Theorem A or the subclass $[(u-\alpha)^2+(v-\beta)^2]$ of Theorem 1.

THEOREM 2. *If $p(u, v)$ is continuous and $\geqq 0$ in D, then $p(u, v)$ is of class PL in D if and only if, for every continuous function $q(u, v)$ of class PL in D,*

$$\frac{1}{\pi\rho^2}\iint_{\xi^2+\eta^2<\rho^2}[p(u_0+\xi, v_0+\eta)][q(u_0+\xi, v_0+\eta)]\,d\xi\,d\eta$$
$$\leqq \left[\frac{1}{2\pi}\int_0^{2\pi}p(u_0+\rho\cos\varphi, v_0+\rho\sin\varphi)\,d\varphi\right] \tag{2}$$
$$\cdot\left[\frac{1}{2\pi}\int_0^{2\pi}q(u_0+\rho\cos\varphi, v_0+\rho\sin\varphi)\,d\varphi\right]$$

holds for every point (u_0, v_0) *in* D *and for every* ρ *such that the circular disc* $(u - u_0)^2 + (v - v_0)^2 \leqq \rho^2$ *is comprised in* D.

Necessity. If $p(u, v)$ and $q(u, v)$ are of class PL in D, then obviously $[p(u, v)][q(u, v)]$ and $[p(u, v)]^{\frac{1}{2}}[q(u, v)]^{\frac{1}{2}}$ are of class PL in D. It follows from Theorem B and the inequality of Schwarz[9] that

$$\frac{1}{\pi\rho^2}\iint_{\xi^2+\eta^2<\rho^2}[p(u_0+\xi, v_0+\eta)][q(u_0+\xi, v_0+\eta)]\,d\xi\,d\eta$$

$$\leqq \left\{\frac{1}{2\pi}\int_0^{2\pi}[p(u_0+\rho\cos\varphi, v_0+\rho\sin\varphi)]^{\frac{1}{2}}[q(u_0+\rho\cos\varphi, v_0+\rho\sin\varphi)]^{\frac{1}{2}}\,d\varphi\right\}^2$$

$$\leqq \left[\frac{1}{2\pi}\int_0^{2\pi}p(u_0+\rho\cos\varphi, v_0+\rho\sin\varphi)\,d\varphi\right]$$

$$\cdot\left[\frac{1}{2\pi}\int_0^{2\pi}q(u_0+\rho\cos\varphi, v_0+\rho\sin\varphi)\,d\varphi\right].$$

Sufficiency. For the given non-negative continuous function $p(u, v)$, we shall use only the fact that (2) holds for the particular functions $q(u, v) = e^{\alpha u+\beta v}$ of class PL. But we could just as well restrict our consideration to the functions $q(u, v) = [(u - \alpha)^2 + (v - \beta)^2]$.

We assume first that $p(u, v)$ is positive and has continuous second derivatives in D. Using finite Taylor expansions, we obtain

$$\frac{1}{\pi\rho^2}\iint_{\xi^2+\eta^2<\rho^2}[p(u_0+\xi, v_0+\eta)][q(u_0+\xi, v_0+\eta)]\,d\xi\,d\eta$$

$$= pq + \tfrac{1}{4}\rho^2[(p_u q_u + p_v q_v) + \tfrac{1}{2}p\Delta q + \tfrac{1}{2}q\Delta p] + o(\rho^2),$$

$$\left[\frac{1}{2\pi}\int_0^{2\pi}p(u_0+\rho\cos\varphi, v_0+\rho\sin\varphi)\,d\varphi\right]\left[\frac{1}{2\pi}\int_0^{2\pi}q(u_0+\rho\cos\varphi, v_0+\rho\sin\varphi)\,d\varphi\right]$$

$$= pq + \tfrac{1}{4}\rho^2[p\Delta q + q\Delta p] + o(\rho^2),$$

where p, q, and their partial derivatives are evaluated at (u_0, v_0), and where $o(\rho^m)$ denotes a quantity (not always the same quantity) such that $o(\rho^m)/\rho^m \to 0$ as $\rho \to 0$. Therefore, by (2),

$$p_u q_u + p_v q_v + \tfrac{1}{2}p\Delta q + \tfrac{1}{2}q\Delta p \leqq p\Delta q + q\Delta p + o(\rho^0).$$

Letting $\rho \to 0$, we obtain

$$p_u q_u + p_v q_v - \tfrac{1}{2}p\Delta q - \tfrac{1}{2}q\Delta p \leqq 0. \tag{3}$$

Putting $q = e^{\alpha u+\beta v}$ in (3), we obtain the quadratic inequality

$$p(\alpha^2 + \beta^2) - 2p_u\alpha - 2p_v\beta + \Delta p \geqq 0$$

[9] See, for instance, Hardy, Littlewood and Pólya, *Inequalities*, Cambridge, 1934, p. 132.

for all α, β. Hence the discriminant satisfies

$$p\Delta p - (p_u^2 + p_v^2) \geqq 0,$$

so that, as in Theorem 1, $p(u, v)$ is of class PL.[10]

We assume now that $p(u, v)$ is only non-negative and has continuous second derivatives in D. Then the function $p(u, v) + \epsilon$ is positive and has continuous second derivatives in D for every $\epsilon > 0$. Further, since $q(u, v)$ is subharmonic, we have

$$\begin{aligned} &\frac{1}{\pi\rho^2}\iint_{\xi^2+\eta^2<\rho^2} q(u_0 + \xi, v_0 + \eta)\, d\xi\, d\eta \\ &\qquad\qquad \leqq \frac{1}{2\pi}\int_0^{2\pi} q(u_0 + \rho\cos\varphi, v_0 + \rho\sin\varphi)\, d\varphi; \end{aligned} \tag{4}$$

it follows from (2) and (4) that (2) is still satisfied if we replace $p(u, v)$ by $p(u, v) + \epsilon$ throughout. Hence, according to the preceding paragraph, $p(u, v) + \epsilon$ is of class PL, and therefore its uniform limit as $\epsilon \to 0$, $p(u, v)$, is of class PL.

Suppose finally that $p(u, v)$ is given to satisfy only the conditions specified in the theorem. That $p(u, v)$ is at least subharmonic follows immediately when we set $\alpha = \beta = 0$ in $q(u, v) = e^{\alpha u + \beta v}$, so that (2) becomes[11]

$$\frac{1}{\pi\rho^2}\iint_{\xi^2+\eta^2<\rho^2} p(u_0 + \xi, v_0 + \eta)\, d\xi\, d\eta \leqq \frac{1}{2\pi}\int_0^{2\pi} p(u_0 + \rho\cos\varphi, v_0 + \rho\sin\varphi)\, d\varphi.$$

For a fixed (u_0, v_0) in D, let the positive quantities ρ and r be so small that the circular disc $(u - u_0)^2 + (v - v_0)^2 \leqq (\rho + r)^2$ is comprised in D. By assumption, we have

$$\begin{aligned} &\frac{1}{\pi\rho^2}\iint_{\xi^2+\eta^2<\rho^2} [p(u_0 + \xi + \xi', v_0 + \eta + \eta')] \\ &\qquad\qquad \cdot[\exp\{\alpha(u_0 + \xi + \xi') + \beta(v_0 + \eta + \eta')\}]\, d\xi\, d\eta \\ &\qquad \leqq \left[\frac{1}{2\pi}\int_0^{2\pi} p(u_0 + \xi' + \rho\cos\varphi, v_0 + \eta' + \rho\sin\varphi)\, d\varphi\right] \\ &\qquad\qquad \cdot\left[\frac{1}{2\pi}\int_0^{2\pi} \exp\{\alpha(u_0 + \xi' + \rho\cos\varphi) + \beta(v_0 + \eta' + \rho\sin\varphi)\}\, d\varphi\right] \end{aligned} \tag{5}$$

[10] If we had chosen the functions $q(u, v) = [(u - \alpha)^2 + (v - \beta)^2]$, the quadratic inequality would have been $[(u_0 - \alpha)^2 + (v_0 - \beta)^2]\Delta p - 4p_u(u_0 - \alpha) - 4p_v(v_0 - \beta) + 4p \geqq 0$, with the same implications as above.

[11] If we had chosen the functions $q(u, v) = [(u - \alpha)^2 + (v - \beta)^2]$, we could show that $p(u, v)$ is subharmonic as we did in the proof of Theorem 1.

for all ξ', η' satisfying $\xi'^2 + \eta'^2 < r^2$, and for every choice of the real constants α, β. Dividing both sides of (5) by the positive constant $e^{\alpha\xi'+\beta\eta'}$, we obtain

$$\frac{1}{\pi\rho^2}\iint_{\xi^2+\eta^2<\rho^2}[p(u_0+\xi+\xi', v_0+\eta+\eta')] \cdot [\exp\{\alpha(u_0+\xi)+\beta(v_0+\eta)\}]\, d\xi\, d\eta$$

$$\leqq \left[\frac{1}{2\pi}\int_0^{2\pi} p(u_0+\xi'+\rho\cos\varphi, v_0+\eta'+\rho\sin\varphi)\, d\varphi\right] \cdot \left[\frac{1}{2\pi}\int_0^{2\pi}\exp\{\alpha(u_0+\rho\cos\varphi)+\beta(v_0+\rho\sin\varphi)\}\, d\varphi\right]. \tag{6}$$

Now integrating both sides of (6) over $\xi'^2 + \eta'^2 < r^2$, dividing by πr^2, and reversing the order of integration, we have

$$\frac{1}{\pi\rho^2}\iint_{\xi^2+\eta^2<\rho^2}[p_r^{(1)}(u_0+\xi, v_0+\eta)][\exp\{\alpha(u_0+\xi)+\beta(v_0+\eta)\}]\, d\xi\, d\eta$$

$$\leqq \left[\frac{1}{2\pi}\int_0^{2\pi} p_r^{(1)}(u_0+\rho\cos\varphi, v_0+\rho\sin\varphi)\, d\varphi\right] \cdot \left[\frac{1}{2\pi}\int_0^{2\pi}\exp\{\alpha(u_0+\rho\cos\varphi)+\beta(v_0+\rho\sin\varphi)\}\, d\varphi\right].$$

That is, (2) is still satisfied for $q(u, v) = e^{\alpha u+\beta v}$ when we replace $p(u, v)$ by $p_r^{(1)}(u, v)$.[12] Similarly, (2) is satisfied by $p_r^{(2)}(u, v)$. Further, $p_r^{(2)}(u, v)$ is non-negative and has continuous second derivatives. As above, then, $p_r^{(2)}(u, v)$ is of class PL. But we have shown that $p(u, v)$ is subharmonic, so that $p_r^{(2)}(u, v) \searrow p(u, v)$ as $r \to 0$. Since the limit of a decreasing sequence of functions of class PL is again of class PL, it follows that $p(u, v)$ is of class PL.

THE UNIVERSITY OF MICHIGAN.

[12] A similar device may be used for the functions $q(u, v) = [(u-\alpha)^2 + (v-\beta)^2]$.

RIEMANNIAN MANIFOLDS WITH POSITIVE MEAN CURVATURE

BY S. B. MYERS

1. Intuitively it is natural to expect that a manifold whose curvature is everywhere positive and bounded away from zero will, if indefinitely extended, be closed. Illustrating this geometric idea, in 1931 Hopf and Rinow[1] proved that a complete surface whose curvature is everywhere greater than or equal to a positive constant e^2 is closed (compact) and has a diameter not exceeding π/e, and its universal covering manifold is also closed with diameter not exceeding π/e. In 1935[2] the present author generalized this result to complete n-dimensional Riemannian manifolds. There the curvature K is a function of a point and a plane of directions; and if $K \geqq e^2$ for all points and all planes of directions, the space is closed and has diameter not exceeding π/e, the same results holding for the universal covering manifold.

The question now arises as to what conclusions can be drawn if the curvature is allowed to vary more freely on a complete manifold, but the mean curvature,[3] a function of a point and a direction, is kept $\geqq e^2$. The result of the present paper is that such a manifold M is closed and has[4] diameter not exceeding a, where

$$a = \frac{\pi\sqrt{n - 1}}{e};$$

also, its universal covering manifold is closed and has diameter not exceeding a. This implies that the fundamental group of M is finite. An important application of this result is to spaces of constant positive mean curvature, which are solutions of the field equations in the general theory of relativity.[5]

2. Let M be an n-dimensional Riemannian manifold[6] of class $C^{(3)}$, whose mean curvature everywhere $\geqq e^2$. We prove first the following lemma.

Received March 3, 1941; presented to the American Mathematical Society, December 31, 1940.

[1] H. Hopf and W. Rinow, *Über den Begriff der vollständigen differential-geometrischen Fläche*, Comm. Math. Helvetici, vol. 3(1931), p. 224.

[2] S. B. Myers, *Riemannian manifolds in the large*, this Journal, vol. 1(1935), pp. 42–43.

[3] The mean curvature at a point P with respect to a vector (v) is the sum of the $n - 1$ 2-dimensional curvatures obtained by pairing (v) with each of a set of $n - 1$ mutually orthogonal vectors all orthogonal to (v), and is independent of the choice of such a set. See Eisenhart, *Riemannian Geometry*, p. 113.

[4] If mean curvature were defined as an average instead of a sum, then this upper bound for diameter would be simply π/e instead of a.

[5] The results of this paper are, however, proved for positive definite metrics, and hence are not immediately applicable to the physical theory.

[6] See, for example, Myers, op. cit., p. 40.

LEMMA. *On an n-dimensional Riemannian manifold M of class $C^{(3)}$ whose mean curvature with respect to every point and every direction $\geqq e^2$, no geodesic arc of length greater than $\pi(n-1)^{\frac{1}{2}}e^{-1}$ can be the shortest arc joining its ends.*

Let g be any geodesic arc of length a, with endpoints P and Q. Set up coördinates $(x) = (x^1, \cdots, x^n)$ in the neighborhood of g such that the functions $g_{\alpha\beta}$ $(\alpha, \beta = 1, \cdots, n)$ appearing in the fundamental quadratic form satisfy the following conditions along g:

$$g_{\alpha\beta} = \delta_{\alpha\beta}, \qquad \frac{\partial g_{\alpha\beta}}{\partial x^\gamma} = 0 \qquad (\alpha, \beta, \gamma = 1, \cdots, n),$$

while the coördinates $x^1, \cdots, x^{n-1}$ are constant along g and x^n is the arc length s on g measured from P.[7]

In this coördinate system the curvature of M at a point $\bar{P}$ on g with respect to the plane of directions determined by the unit vector $(0, 0, \cdots, 1)$ tangent to g at $\bar{P}$ and any vector $(\eta^1, \eta^2, \cdots, \eta^{n-1}, 0)$ orthogonal to g at $\bar{P}$ is given by[8]

$$K(\bar{P}, \eta) = \frac{R_{nink}\eta^i\eta^k}{\eta^j\eta^j} \qquad (i, j, k = 1, \cdots, n-1).$$

The functions $R_{nink}(s)$ are Riemann symbols of the first kind in the coördinates (x) taken along g. The mean curvature $\bar{K}$ at $\bar{P}$ with respect to the vector $(0, 0, \cdots, 1)$ is given by

$$\bar{K}(\bar{P}) = R_{nknk} = -R_{nn},$$

which is the trace of the determinant $| R_{nink} |$.

Now consider the second variation $J(\eta)$ along g in the non-parametric calculus of variations problem associated with

$$\int_0^a \left(g_{\alpha\beta} \frac{dx^\alpha}{dx^n} \frac{dx^\beta}{dx^n} \right)^{\frac{1}{2}} dx^n.$$

$J(\eta)$ is to be evaluated for arbitrary functions $\eta^1(s), \cdots, \eta^{n-1}(s)$ of class D' vanishing at $s = 0$ and $s = a$, and must be positive for all such (η) not $\equiv (0)$ unless g contains a point conjugate to P (perhaps at Q). Upon computation, we find[9]

$$J(\eta) = \int_0^a (\eta'^i\eta'^i - R_{nink}\eta^i\eta^k)\, ds.$$

[7] See T. Levi-Civita, Math. Annalen, vol. 97(1927), pp. 291–320. Also see M. Morse, *The Calculus of Variations in the Large*, Am. Math. Soc. Colloquium Publication, New York, 1934, pp. 108–110.

[8] We are using the summation convention of tensor analysis.

[9] Cf. J. L. Synge, *On the neighborhood of a geodesic in Riemannian space*, this Journal, vol. 1(1935), p. 527.

Integrating the first term by parts, and using the fact that (η) vanishes at both ends, we obtain

$$J(\eta) = -\int_0^a (\eta^i \eta^{i\prime\prime} + R_{nink}\eta^i \eta^k)\, ds$$

$$= -\int_0^a [\eta^i \eta^{i\prime\prime} + K(s, \eta)\eta^i \eta^i]\, ds.$$

Now $\eta^i(s)$ can be written as $f(s)\zeta^i(s)$, where

$$\zeta^i\zeta^i \equiv 1, \qquad \zeta^i\zeta^{i\prime} = 0, \qquad \zeta^i\zeta^{i\prime\prime} = -\zeta^{i\prime}\zeta^{i\prime}$$

and

$$f(0) = f(a) = 0.$$

We find

$$J(\eta) = -\int_0^a [f\zeta^i(f\zeta^{i\prime\prime} + 2f'\zeta^{i\prime} + f''\zeta^i) + Kf^2]\, ds$$

$$= -\int_0^a f[f'' + f(K - \zeta^{i\prime}\zeta^{i\prime})]\, ds.$$

$J(\eta)$ must be greater than 0 for $f(s) \not\equiv (0)$ vanishing at both ends and all unit vectors $\zeta^i(s)$ unless g contains a point conjugate to P.

Now choose

$$f(s) = \sin e(n-1)^{-\frac{1}{2}}\, s,$$

so that

$$f'' + \frac{e^2}{n-1} f = 0,$$

and choose $n - 1$ constant unit vectors $\zeta_j^i(s) \equiv \delta_j^i$. For $\eta_j^i = f\zeta_j^i$ we obtain

$$J(\eta_j) = -\int_0^a f[f'' + fK(s, \eta_j)]\, ds = \int_0^a f^2\left(\frac{e^2}{n-1} - K\right) ds$$

$$\sum_{j=1}^{n-1} J(\eta_j) = \int_0^a f^2(e^2 - \bar{K})\, ds.$$

But $\bar{K} \geqq e^2$, so that

$$\sum_{j=1}^{n-1} J(\eta_j) \leqq 0.$$

Hence for at least one value of j, $J(\eta_j) \leqq 0$.

Therefore g must contain a point conjugate to P (perhaps at Q). Thus no geodesic arc of length greater than a can be a minimizing arc, and the lemma is proved.

3. Now suppose M is complete. Then[10] every pair of points P and Q on M can be joined by a shortest curve which will be a geodesic with length equal to the distance between its ends. From the lemma it follows that the shortest curve joining P to Q is of length not exceeding a, hence the distance $PQ \leqq a$. This makes every set of points on M bounded. But completeness implies that every infinite bounded set on M has a limit point,[10] so that every infinite set on M has a limit point and M is compact (closed).

We have, then, the following:

THEOREM I. *A complete n-dimensional Riemannian manifold whose mean curvature at every point and with respect to every direction* $\geqq e^2$ *is closed and has diameter not exceeding* $\pi(n-1)^{\frac{1}{2}}e^{-1}$.

Now the universal covering manifold of M can be provided with the same local differential geometry as M, and will also be complete. So we have

THEOREM II. *The universal covering manifold of a complete n-dimensional Riemannian manifold with mean curvature everywhere not less than* e^2 *is closed and has diameter not exceeding* $\pi(n-1)^{\frac{1}{2}}e^{-1}$.

COROLLARY. *A complete n-dimensional Riemannian manifold with mean curvature everywhere at least equal to* e^2 *has a finite fundamental group.*

As a special case of these theorems we have the

COROLLARY. *A complete n-dimensional Riemannian manifold with constant positive mean curvature* e^2 *is closed, has a finite fundamental group, and has diameter not exceeding* $\pi(n-1)^{\frac{1}{2}}e^{-1}$.

4. The n-dimensional unit sphere, which has constant unit curvature and hence constant mean curvature $= n - 1$, and has diameter π, is an example showing that the upper bound for diameter just given cannot be improved. An example to show that the results of this paper go beyond the previously treated case of manifolds with positive curvature bounded away from zero is the topological product of two 2-dimensional unit spheres, provided with a local Riemannian quadratic form which is the sum of the forms for the two 2-spheres. This 4-dimensional manifold has constant positive mean curvature 1. The curvature itself varies from 0 to 1 inclusive, and thus is not constant and not even bounded away from 0.

UNIVERSITY OF MICHIGAN.

[10] Hopf and Rinow, loc. cit.

AN ANALOGUE OF THE BERNOULLI POLYNOMIALS

By L. Carlitz

1. We define a set of polynomials $\beta_m(u)$ with coefficients in $GF(p^n, x)$ by means of[1]

$$\frac{\psi(tu)}{u\psi(t)} = \sum_{m=0}^{\infty} \frac{\beta_m(u)}{g_m} t^m. \tag{1.1}$$

The function $\psi(t)$ may be defined by[2]

$$\psi(t) = \sum_{i=0}^{\infty} (-1)^i \frac{t^{p^{ni}}}{F_i}, \tag{1.2}$$

where

$$F_k = (x^{p^{nk}} - x)(x^{p^{nk}} - x^{p^n}) \cdots (x^{p^{nk}} - x^{p^{n(k-1)}}), \qquad F_0 = 1;$$

and the denominator in the right member of (1.1) is given by

$$g_m = g(m) = F_0^{a_0} F_1^{a_1} \cdots F_s^{a_s}, \tag{1.3}$$

where

$$m = a_0 + a_1 p^n + \cdots + a_s p^{ns} \qquad (0 \leqq a_i < p^n).$$

Thus it is clear that the coefficients in $\beta_m(u)$ are rational functions over the $GF(p^n)$, indeed over the $GF(p)$.

If we put

$$\frac{t}{\psi(t)} = \sum_{m=0}^{\infty} \frac{B_m}{g_m} t^m, \tag{1.4}$$

then by (1.1) and (1.2)

$$\beta_m(u) = \sum_{p^{ni} \leqq m+1} (-1)^i \frac{g(m)}{F_i g(m - p^{ni} + 1)} B_{m-p^{ni}+1} u^{p^{ni}-1}. \tag{1.5}$$

On the other hand, if we use the expansion

$$\psi(tu) = \sum_{i=0}^{\infty} (-1)^i \frac{\psi_i(u)}{F_i} \psi^{p^{ni}}(t), \tag{1.6}$$

Received April 10, 1941.

[1] The right member of (1.1) contains only terms for which $p^n - 1 \mid m$; $\beta_m(u)$ is defined for such m only.

[2] For the properties of $\psi(t)$ and $\psi_k(t)$ used here see L. Carlitz, *On certain functions connected with polynomials in a Galois field*, this Journal, vol. 1(1935), pp. 137–168.

then we get

$$\beta_m(u) = \frac{1}{u} \sum_{p^{ni} \leq m+1} (-1)^i \frac{\psi_i(u)}{F_i} A_m^{(i)}, \tag{1.7}$$

where $A_m^{(i)}$ is given by

$$\psi^{p^{ni}-1}(t) = \sum_{m=0}^{\infty} \frac{1}{g_m} A_m^{(i)} t^m, \tag{1.8}$$

and

$$\psi_k(u) = \sum_{i=0}^{k} (-1)^{k-i} \frac{F_k}{F_i L_{k-i}^{p^{ni}}} u^{p^{ni}},$$

where

$$L_k = (x^{p^{nk}} - x)(x^{p^{n(k-1)}} - x) \cdots (x^{p^n} - x), \qquad L_0 = 1.$$

We also remark that $\psi_k(u) = \prod (u + M)$ the product extending over all M (including 0) of degree $< k$.

In (1.1) the value $u = 0$ is excluded; however, it is clear from (1.5) that

$$\beta_m(0) = B_m .$$

Also from (1.7) follows

$$\beta_m(1) = 0 \qquad \text{for } m > 0, \tag{1.9}$$

since $\psi_m(1) = 0$ for $m > 0$. We may also notice that

$$\beta_m(cu) = \beta_m(u) \qquad (c \text{ in } GF(p^n),\ c \neq 0). \tag{1.10}$$

(1.9) and (1.5) yield a recursion formula for B_m. More generally, from (1.1) we get the recursion for $\beta_m(u)$:

$$\sum_{p^{ni} \leq m} \frac{(-1)^i}{F_i g(m - p^{ni})} \beta_{m-p^{ni}}(u) = \begin{cases} \dfrac{(-1)^k}{F_k} u^{p^{nk}-1} & \text{for } m = p^{nk}, \\ 0 & \text{otherwise.} \end{cases}$$

Properties of B_m have already been discussed.[3] In the present note we derive various theorems on the arithmetic nature of $\beta_m(u)$, when u is restricted to polynomial values. Since the methods are similar to those of the earlier papers, we shall not give the proofs in detail.

2. In $\beta_m(u)$ we now put $u = U$, a polynomial in x; in view of (1.10) we may assume that U is primary. By a general theorem[4] $A_m^{(i)}$ is integral, that is, a

[3] L. Carlitz, *An analogue of the von Staudt-Clausen theorem*, this Journal, vol. 3(1937), pp. 503–517, and *An analogue of the Staudt-Clausen theorem*, this Journal, vol. 7(1940), pp. 62–67. These papers will be cited as I and II, respectively.

[4] I, p. 507.

polynomial in x; also it is easy to show that $\psi_i(U)/F_i$ is integral for all U. Hence applying (1.6) we get at once

THEOREM 1. *For all polynomials $U = U(x)$ and all m (such that $p^n - 1 \mid m$), the product $U\beta_m(U)$ is integral.*

Thus the denominator of $\beta_m(U)$ contains only factors of U. We can easily improve this result. From the identity

$$\psi_k(u) = \psi_{k-1}^{p^n}(u) - F_{k-1}^{p^n-1}\psi_{k-1}(u)$$

follows

$$\frac{\psi_k(U)}{Ug(p^{nk}-1)} = \left\{\frac{\psi_{k-1}^{p^n-1}(U)}{F_{k-1}^{p^n-1}} - 1\right\}\frac{\psi_{k-1}(U)}{Ug(p^{n(k-1)}-1)},$$

and therefore

$$\frac{\psi_k(U)}{Ug(p^{nk}-1)}$$

is integral.[5] Note that by (1.3)

$$g(p^{nk}-1) = (F_0F_1\cdots F_{k-1})^{p^n-1} = \frac{F_k}{L_k}.$$

We require also the fact that[6]

$$g(p^{nk}-1) \mid A_m^{(k)}. \tag{2.1}$$

Hence (1.7) may be put in the form

$$\beta_m(U) = \sum_k \frac{F_k}{L_k^2} G_k, \tag{2.2}$$

where G_i is integral. Now if the fraction F_kG_k/L_k^2 is put in reduced form, it is clear that the irreducible factors (if any) of the denominator are simple, and except possibly for the case $p^n = k = 2$, are of degree k. For $p^n = k = 2$, F_k/L_k^2 reduces to $1/(x^4 - x)$, and therefore the denominator of F_2G_2/L_2^2 may involve factors of the first degree also. This proves

THEOREM 2. *If we put $\beta_m(U) = N/D$, where N and D are relatively prime, then D has simple factors only.*

By Theorem 1 it follows that D consists of a product of some (possibly none) of the irreducible divisors of U. We call $D = D_m(U)$ the denominator of $\beta_m(U)$.

In the next place for G an arbitrary polynomial, consider

$$\sum_m \frac{G^m - 1}{g_m}\beta_m(U)t^m = \frac{\psi(tGU)}{U\psi(tG)} - \frac{\psi(tU)}{U\psi(t)}. \tag{2.3}$$

[5] This is a special case of a general theorem. See Carlitz, *A set of polynomials*, this Journal, vol. 6(1940), pp. 486–504; p. 502, Lemma 2.

[6] I, Theorem 3.

Now in (1.6) put $u = G$ and substitute in the right member of (2.3). We may prove[7]

THEOREM 3. *For all G the product*

$$G(G^m - 1)\beta_m(U) \tag{2.4}$$

is integral.

Suppose next that P is any irreducible factor of the denominator of $\beta_m(U)$, that is, $P \mid D$ in the notation of Theorem 2. Since in (2.4) G is arbitrary, we may take it equal to a primitive root (mod P); it follows that m is a multiple of $p^{nk} - 1$, where k is the degree of P.

THEOREM 4. *If, in the notation of Theorem 2, P is an irreducible divisor of D of degree k, then $p^{nk} - 1$ must divide m.*

As a consequence of Theorems 2 and 4 we may put (1.7) in the following form

$$\beta_m(U) = G_m + \frac{1}{U} \sum_{p^{nk}-1 \mid m} (-1)^k \frac{\psi_k(U)}{F_k} A_m^{(k)}, \tag{2.5}$$

where G_m is integral and the summation extends only over such k for which $p^{nk} - 1 \mid m$.

3. We now make use of the following result:[8]

$$\psi^{p^{nk}-1} \equiv \left(\sum_{i=0}^{\infty} \frac{(-1)^{ki}}{F_{ki}} t^{p^{nki}} \right)^{p^{nk}-1} \pmod{P}, \tag{3.1}$$

where P is irreducible of degree k. From (3.1) and (1.8) it follows that

$$A_m^{(k)} \equiv \frac{(-1)^{k-1+\delta k+nk}}{(\delta_h!)} \pmod{P}, \tag{3.2}$$

where

$$m = \sum_h \delta_h p^h \qquad (0 \leqq \delta_h < p) \tag{3.3}$$

and

$$\delta \equiv \sum_{i,j} i\delta_{nki+j} \pmod{2}; \tag{3.4}$$

also the conditions

$$nk(p-1) = \sum_h \delta_h, \qquad p^{nk} - 1 \mid m \tag{3.5}$$

must hold. Note that, for fixed m, (3.5) is satisfied by at most one value of k. If (3.5) does not hold, then $A_m^{(k)} \equiv 0 \pmod{P}$.

[7] For the proof compare I, Theorem 5.

[8] I, Theorem 8; for a simplified proof see II, p. 66.

4. Returning to (2.5), from the above we see that

$$\beta_m(U) = G_m + (-1)^k \frac{\psi_k(U)}{U \cdot F_k} A_m^{(k)} \qquad (p^n \neq 2), \tag{4.1}$$

where k satisfies (3.5); if, however, (3.5) does not hold, then $\beta_m(U)$ is integral. To simplify the right member of (4.1) note that if P is irreducible of degree k, then

$$\frac{\psi_k(U)}{U} = \prod_{\substack{\deg M < k \\ M \neq 0}} (U + M) \equiv \begin{cases} -1 & \text{for } P \mid U, \\ 0 & \text{for } P \nmid U, \end{cases}$$

the congruences being (mod P). If we now apply (3.2), we get the following

THEOREM 5. ($p^n \neq 2$.) *Let $p^n - 1 \mid m$; put*

$$m = \sum_h \delta_h p^h \qquad (0 \leqq \delta_h < p).$$

If the system[9]

$$nk(p-1) = \sum_h \delta_h, \qquad p^{nk} - 1 \mid m \tag{4.2}$$

is inconsistent, then $\beta_m(U)$ is integral; if (4.2) is consistent, then k is uniquely determined and

$$\beta_m(U) = G_m - e \sum_{P \mid U} \frac{1}{P}, \tag{4.3}$$

where G_m is integral, the summation is extended over those irreducible P of degree k that divide U and

$$e = \frac{(-1)^{nk+\delta k}}{\prod_h (\delta_h!)}, \qquad \delta \equiv \sum_{i,j} i\delta_{nki+j} \qquad (\text{mod } 2).$$

For $p^n = 2$, there is the possibility that the denominator may contain additional irreducibles of the first degree. Note that

$$\frac{\psi_2(U)}{U} = (U+1)(U+x)(U+x+1)$$

$$\equiv \begin{cases} 0 \quad (\text{mod } x^2) & \text{for } x^2 \nmid U, \\ x \quad (\text{mod } x^2) & \text{for } x^2 \mid U, \end{cases}$$

with similar results (mod $(x+1)^2$). A detailed examination of $A_m^{(2)}$ now leads to[10]

[9] Compare II, p. 63.

[10] Compare I, Theorem 10. Note that the last part of that theorem is incorrectly stated. It should read: "If (7.3) is inconsistent, then for m odd of the form $2^a + 1$,

$$B_m = G_m + \frac{1}{x} + \frac{1}{x+1},$$

while for other m, B_m is integral."

THEOREM 6. ($p^n = 2$.) *If the system* (4.2) *is consistent for* $k \neq 2$, *then*

$$\beta_m(U) = G_m + \sum_{\substack{\deg P = k \\ P \mid U}} \frac{1}{P}; \tag{4.4}$$

if $k = 2$, *m even, then*

$$\beta_m(U) = G_m + \frac{\epsilon}{x^2 + x + 1}, \tag{4.5}$$

where $\epsilon = 1$ *for* $x^2 + x + 1 \mid U$, $\epsilon = 0$ *for* $x^2 + x + 1 \nmid U$; *while if* $k = 2$, *m odd, then*

$$\beta_m(U) = G_m + \frac{\epsilon}{x^2 + x + 1} + \sum_{\substack{\deg Q = 1 \\ Q^2 \mid U}} \frac{1}{Q}. \tag{4.6}$$

If (4.2) *is inconsistent, then for m odd of the form* $2^a + 1$,

$$\beta_m(U) = G_m + \sum_{\substack{\deg Q = 1 \\ Q^2 \mid U}} \frac{1}{Q}, \tag{4.7}$$

while for other m, $\beta_m(U)$ *is integral.*

As an immediate corollary of Theorem 5 we may state

THEOREM 7. ($p^n \neq 2$.) *For U and V relatively prime,*

$$\beta_m(UV) = \beta_m(U) + \beta_m(V) + G_m .$$

In the next place it is clear that (for $p^n \neq 2$)

$$\beta_m(P^i) = \beta_m(P^{i+1}) + G_m \qquad (i > 0),$$

and this leads to the following

THEOREM 8. ($p^n \neq 2$.) *For arbitrary U,*

$$\beta_m(U) = G_m + \sum_{P \mid U} \beta_m(P),$$

the summation extending over all irreducible P dividing U.

We remark that the last two theorems are not true for $p^n = 2$, as is clear from an examination of the several cases of Theorem 6. The case of failure is that arising from (4.7); in all other cases the theorems hold.

Comparison of Theorem 5 with the corresponding result[11] concerning B_m leads to

THEOREM 9. ($p^n \neq 2$.) *If U is a multiple of* D_m, *the denominator of* B_m, *then* $\beta_m(U) - \beta_m$ *is integral.*

In particular, in view of (2.3), we may take $U = G(G^m - 1)$, where G is

[11] I, Theorem 9.

arbitrary; the simplest choice is $U = x(x^m - 1)$. On the other hand if (4.2) holds, we may take for U any multiple of

$$\Theta_k = \prod_{\deg P = k} P;$$

a particularly simple choice is $U = x^{p^{nk}} - x$. Note again that the last theorem may fail for $p^n = 2$.

5. Results of a different nature may also be derived from (1.7). Define the operator Δ^i by means of

$$\Delta f(u) = f(xu) - xf(u),$$
$$\Delta^{k+1} f(u) = \Delta^k f(xu) - x^{p^{nk}} \Delta^k f(u).$$

It may be verified that for $k \leqq i$

$$\Delta^k \psi_i(u) = \frac{F_i}{F_{i-k}^{p^{nk}}} \psi_{i-k}^{p^{nk}}(u).$$

Thus (1.7) yields

$$\Delta^k \{U\beta_m(U)\} = \sum_i (-1)^{i+k} \left\{\frac{\psi_i(U)}{F_i}\right\}^{p^{nk}} A_m^{(i+k)}. \tag{5.1}$$

Clearly by (2.1) each term on the right side of (5.1) is a multiple of $g(p^{nk} - 1)$, and therefore

$$\Delta^k \{U\beta_m(U)\} \equiv 0 \qquad \left(\mathrm{mod}\ \frac{F_k}{L_k}\right). \tag{5.2}$$

For $U = 1$, (5.1) becomes

$$\Delta^k \{u\beta_m(u)\}_{u=1} = (-1)^k A_m^{(k)}, \tag{5.3}$$

which incidentally expresses $A_m^{(k)}$ in terms of $\beta_m(x^i)$.

It is easy to derive another formula of this type from (1.7) in the following way. In (1.7) replace u by uM, and let M run through the primary polynomials of degree k. Since[12]

$$\sum_{\deg M = k} \frac{\psi_i(Mu)}{M} = (-1)^k \frac{F_i}{L_k F_{i-k}^{p^{nk}}} \psi_{i-k}^{p^{nk}}(u),$$

we get

$$\sum_{\deg M = k} \beta_m(Mu) = \frac{1}{u} \sum_i \frac{(-1)^i}{L_k} \left\{\frac{\psi_i(u)}{F_i}\right\}^{p^{nk}} A_m^{(i+k)}, \tag{5.4}$$

which may be compared with (5.1).

[12] See L. Carlitz, *Some sums involving polynomials in a Galois field*, this Journal, vol. 5(1939), pp. 941–947; p. 943, formula (4.4).

In (5.4) put $u = 1$:

$$\sum_M \beta_m(M) = \frac{1}{L_k} A_m^{(k)},$$

which is similar to (5.3). Comparison with the results of §3 now leads to the following

THEOREM 10. ($p^n \neq 2$.) *The sum*

$$\sum_{\deg M = k} \beta_m(M) \tag{5.5}$$

is integral if m *and* k *do not satisfy* (4.2); *if, however,* (4.2) *holds, then the sum* (5.5) *equals*

$$G_m - e \sum_{\deg P = k} \frac{1}{P},$$

where e *has the same meaning as in Theorem* 5.

As a corollary it follows that

$$\sum_{\deg M = k} \beta_m(M) = B_m + G_m.$$

In this connection we may note that (1.5) yields

$$\sum_{\deg M < s}{}' \beta_m(M) = B_m,$$

the summation extending over all primary M of degree $< s$, where $p^{ns} > m + 1$.

DUKE UNIVERSITY.

A MINIMUM PROBLEM IN THE THEORY OF ANALYTIC FUNCTIONS

By J. L. Doob

Let $f(z)$ be a complex-valued function defined on $\gamma: |z| = 1$. F. Riesz[1] has investigated the problem of minimizing the integral

$$\int_\gamma |f(z) - p(z)|\,|dz| \tag{0.1}$$

for functions $p(z)$ which are the boundary functions on γ of functions analytic within γ. Riesz assumed $f(z)$ to be a polynomial in $1/z$. He found that there is a minimizing function $p(z) = g(z)$ (uniquely determined in the sense that two minimizing functions differ at most on a set of measure 0), which is a polynomial. The zeros of $f(z) - g(z)$ are distributed in a simple way, and Riesz' uniqueness proof depends on that fact. Riesz' problem can also be considered as the problem of minimizing $\int_\gamma |h(z)|\,|dz|$ for functions $h(z)$ which are the boundary values of functions $h(z)$ analytic within γ, a finite number of whose initial power series coefficients are given. In this form, the problem has been generalized by Kakeya,[2] who imposes other conditions on the functions $h(z)$ within γ: the values of $h(z)$ and its derivatives are prescribed at given points. In all this work, the uniqueness theorem depends essentially on the fact that the minimizing $h(z)$ can be written down explicitly, or at least that the minimizing $h(z)$ is rational with a known peculiar distribution of zeros and poles.

Before describing the problem which is to be solved in this paper, we shall state the facts which will be needed from the theory of analytic functions.[3] Let $f(z)$ be a complex-valued Lebesgue integrable function, defined on γ. Then $f(z)$ has a Fourier series:

$$f(z) \sim \sum_{-\infty}^{\infty} a_n z^n, \qquad a_n = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z^{n+1}}\,dz = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})e^{-ni\theta}\,d\theta. \tag{0.2}$$

The function $f(z)$ is said to be of power series type if $a_{-1} = a_{-2} = \cdots = 0$. If $f(z)$ is of power series type, $\sum_0^\infty a_n z^n$ converges within γ to an analytic function

Received March 24, 1941.

[1] F. Riesz, *Über Potenzreihen mit vorgeschriebenen Anfangsgliedern*, Acta Mathematica, vol. 42(1920), pp. 145–171.

[2] S. Kakeya, Proceedings of the Physico-Mathematical Society of Japan, (3), vol. 3(1921), pp. 48–58.

[3] See for the results summarized in this paragraph, F. Riesz, *Ueber die Randwerte einer analytischen Funktion*, Mathematische Zeitschrift, vol. 18(1923), pp. 87–95; F. and M. Riesz, *Ueber die Randwerte einer analytischen Funktion*, Compte Rendu du Quatrième Congrès des Mathématiciens Scandinaves à Stockholm (1916), pp. 27–44.

$f_1(z)$, since $a_n \to 0$. Moreover, $f_1(z)$ within γ is given by the Cauchy integral formula, if the boundary function $f(z)$ is used, and

$$\lim_{r \uparrow 1} f_1(re^{i\theta}) = f_1(e^{i\theta}) \tag{0.3}$$

exists for almost all θ, with $f_1 = f$ for almost all θ. Thus $f(z)$ coincides almost everywhere on γ with the boundary function of an analytic function. A boundary function $f_1(z)$ has many properties which we shall use below without further reference. For example: its discontinuities, if any, cannot be jumps; if it is of bounded variation, it must be absolutely continuous; if it vanishes on a set of positive measure, it vanishes identically. We shall also use the fact that if $f(z)$ and $g(z)$ are of power series type, any linear combination is of power series type, and if $g(z)$ is bounded, $f \cdot g$ is of power series type.[4]

The purpose of the present paper is to minimize the integral (0.1) for any given $f(z)$ integrable on γ, where $p(z)$ ranges through the functions of power series type. It will be shown that there is a minimizing $p(z) = g(z)$, uniquely determined (neglecting sets of measure 0), and the properties of $f(z) - g(z)$ will be investigated. The class of minimal functions $f(z) - g(z)$ will be completely described.

The following facts on functions of power series type will be useful. The proofs are obvious.

LEMMA 1. *If $f(z)$ and $z^N\overline{f(z)}$ are both of power series type, then if $N < 0$, $f(z) = 0$ almost everywhere on γ, and if $N \geqq 0$, $f(z)$ coincides almost everywhere on γ with a polynomial of degree not exceeding N (or the null polynomial). If $R(z)$ is real, and if $z^N R(z)$ is of power series type, then if $N < 0$, $z^N R(z) = 0$ almost everywhere on γ, and if $N \geqq 0$, $z^N R(z)$ coincides almost everywhere on γ with a polynomial of degree not exceeding $2N$ (or the null polynomial).*

We shall need the following theorem, generalizing a result proved by Riesz.[5]

THEOREM 2. *Let $f(z)$, $g(z)$ be Lebesgue integrable functions defined on $\gamma: |z| = 1$, where $|g| > 0$ on a set of positive measure. Define $\xi(\lambda)$ by*

$$\xi(\lambda) = \int_\gamma |f(z) + \lambda g(z)| \, |dz|. \tag{2.1}$$

Then $\xi(\lambda)$ is a continuous function of the complex parameter λ, and $\xi(\lambda) \to \infty$ when $|\lambda| \to \infty$. The function $\xi(\lambda)$ assumes its minimum value ξ_0 either at a single point or on a closed (finite) line segment. Moreover, $\xi(0) = \xi_0$ if and only if

$$\left| \int_{f \neq 0} g \frac{|f|}{f} |dz| \right| \leqq \int_{f=0} |g| \, |dz|, \tag{2.2}$$

[4] Op. cit. (footnote 1), pp. 161–162.
[5] Op. cit. (footnote 1), pp. 159–160.

and $\xi(\lambda) = \xi_0$ *on the real interval* $(-1, 1)$ *if and only if* $g(z) = \sigma(z)f(z)$ *almost everywhere on* γ, *where* $\sigma(z)$ *is real-valued,* $|\sigma| \leqq 1$, *and*

$$(2.2') \qquad \int_\gamma \sigma(z) \mid f(z) \mid \mid dz \mid = 0.$$

The function $\xi(\lambda)$ is evidently continuous, and becomes infinite when $|\lambda|$ does. If α, β are positive numbers with sum 1,

$$(2.3) \qquad \xi(\alpha\lambda_1 + \beta\lambda_2) \leqq \alpha\xi(\lambda_1) + \beta\xi(\lambda_2),$$

so $\xi(\lambda)$ assumes its minimum value ξ_0 on a closed bounded convex point set. If $\lambda = \rho e^{i\theta}$, where ρ, θ are real, $\xi(\lambda)$ (θ fixed) assumes its minimum as ρ varies either at a single point or on an interval of ρ-values. It is easily verified[6] that

$$(2.4) \qquad \lim_{\rho \downarrow 0} \frac{\xi(\rho e^{i\theta}) - \xi(0)}{\rho} = \Re\left\{e^{i\theta} \int_{f\neq 0} g \frac{|f|}{f} \mid dz \mid\right\} + \int_{f=0} \mid g \mid \mid dz \mid,$$

where $\Re(w)$ denotes the real part of the complex number w. If $\xi(0) = \xi_0$, the difference quotient must be not less than 0 for every θ. Then (2.2) is true. Conversely, if (2.2) is true, the derivative in (2.4) is not less than 0 for any θ, and then, because of (2.3), $\xi(0) = \xi_0$. Now suppose that $\xi(\lambda) = \xi_0$ on the (real) interval $(-1, 1)$. There is equality in

$$(2.5) \qquad |f| \leqq \tfrac{1}{2}|f + g| + \tfrac{1}{2}|f - g|$$

only if either $\arg(f + g) = \arg(f - g)$ or $f = \pm g$, that is only if $g = \sigma f$, where σ is real and $|\sigma| \leqq 1$. But the integral over γ of both sides of (2.5) is ξ_0, so $g = \sigma(z)f$ almost everywhere on γ, where $\sigma(z)$ is real, $|\sigma(z)| \leqq 1$. In this case (2.2) becomes (2.2′). Conversely, if $g = \sigma f$ almost everywhere on γ, where σ is as described above, and if (2.2′) is true, then (2.2) is true, so that $\xi(0) = \xi_0$, and it is easily verified that $\xi(\lambda) = \xi(0) = \xi_0$ on the interval $(-1, 1)$. To finish the proof of the theorem we must show that the equation $\xi(\lambda) = \xi_0$ cannot be true on an open set. It is no restriction to assume that the open set includes a circle of radius 1 with the origin as center. Then if θ is any real number,

$$\int_\gamma |f + \lambda e^{i\theta} g| \mid dz \mid$$

assumes its minimum value on the real interval $(-1, 1)$. But then $e^{i\theta}g = \sigma_\theta(z)f$ almost everywhere on γ, where $\sigma_\theta(z)$ is real, and this is certainly impossible.

With these preparatory results established, we can now treat the minimum problem which is the subject of this paper.

[6] Cf. F. Riesz, op. cit. (footnote 1), p. 160, footnote.

THEOREM 3. *Let $f(z)$ be a Lebesgue integrable function, defined on $\gamma:|z| = 1$. There is then one and only one function $g(z)$ (neglecting differences on sets of measure 0) of power series type such that*

$$(3.1) \qquad \int_\gamma |f(z) - g(z)|\,|dz| \leqq \int_\gamma |f(z) - p(z)|\,dz,$$

if $p(z)$ is of power series type.

Although the proof of the existence of a minimizing function $g(z)$ follows Riesz, we shall outline it, since the same type of argument will be needed again. Let ξ_0 be the greatest lower bound of the integrals on the right in (3.1), for $p(z)$ of power series type. Choose $p_n(z)$ of power series type so that

$$(3.2) \qquad \lim_{n\to\infty} \int_\gamma |f(z) - p_n(z)|\,|dz| = \xi_0.$$

Define $F(z)$, $P_n(z)$ by

$$(3.3) \qquad F(z) = \int_1^z f(z)\,dz, \qquad P_n(z) = \int_1^z p_n(z)\,dz$$

(where the integration is over the arc $(1, z)$ of γ in a counter-clockwise sense). Since $p_n(z)$ is of power series type, there is only one value, 0, assigned to $P_n(1)$, but $F(1)$ may have two values, corresponding to the null arc on γ, and the full perimeter. Then the sequence $\{P_n(z)\}$ is a uniformly bounded sequence of functions of power series type, of uniformly bounded variation, at most

$$\xi_0 + 1 + \int_\gamma |f(z)|\,|dz|$$

for large n. By Helly's theorem, some subsequence converges everywhere on γ. We can suppose that the original sequence has been chosen to be convergent: $P_n(z) \to \tilde{G}(z)$. Then $F(z) - P_n(z) \to F(z) - \tilde{G}(z)$, so

$$(3.4) \quad \text{Var.}\,(F - \tilde{G}) \leqq \liminf_{n\to\infty} \text{Var.}\,(F - P_n) = \liminf_{n\to\infty} \int_\gamma |f - p_n|\,|dz| = \xi_0.$$

Moreover, $\tilde{G}(z)$ is of power series type, and of bounded variation, so as we have remarked above, it coincides almost everywhere (and therefore everywhere except at possible discontinuities) with an absolutely continuous function $G(z)$. Evidently $G'(z) = g(z)$ is of power series type, and

$$(3.5) \qquad \text{Var.}\,(F - G) = \int_\gamma |f - g|\,|dz| \leqq \text{Var.}\,(F - \tilde{G}) \leqq \xi_0.$$

Then by the definition of ξ_0, there must be equality throughout (3.5), and $g(z)$ is the required minimizing function.

Now suppose that $g_1(z)$, $g_2(z)$ are minimizing functions which differ on a set of positive measure. We shall derive a contradiction from this hypothesis.

The integral

$$\int_\gamma |f - g_1 + \lambda(g_1 - g_2)| \, |dz| \tag{3.6}$$

takes on its minimum value ξ_0 at $\lambda = 0$ and $\lambda = 1$. If $\varphi = f - \frac{1}{2}(g_1 + g_2)$, $\psi = \frac{1}{2}(g_2 - g_1)$, then $|\psi| > 0$ on a set of positive measure, and the integral

$$\int_\gamma |\varphi + \lambda\psi| \, |dz| \tag{3.7}$$

takes on its minimum value ξ_0 when $\lambda = \pm 1$. By Theorem 2, $\xi(\lambda) = \xi_0$ on the real interval $(-1, 1)$, and $\psi = \sigma\varphi$ almost everywhere on γ, where σ is real and $|\sigma| \leqq 1$. Moreover, since ψ is of power series type, and $|\psi| > 0$ on a set of positive measure, ψ, and therefore φ, can vanish almost nowhere on γ. By definition of ξ_0, if $n \geqq 0$,

$$\int_\gamma |\varphi + \lambda\psi z^n| \, |dz| \tag{3.8}$$

takes on its minimum value ξ_0 when $\lambda = 0$. Then, by Theorem 2,

$$\int_\gamma \frac{|\varphi|}{\varphi} \psi z^n \, |dz| = 0. \tag{3.9}$$

This means that $|\varphi| \psi/(z\varphi) = p(z)$ is of power series type. But then

$$\pm |\psi| = \sigma |\varphi| = zp(z)$$

is also of power series type. According to Lemma 1, this implies that $\psi = 0$, almost everywhere on γ, contrary to hypothesis.

If $f(z)$ is a given function, integrable on γ, we have shown that there is a corresponding function $g(z)$ of power series type, uniquely determined, neglecting sets of measure 0, minimizing $\int_\gamma |f - p| \, |dz|$ for all functions $p(z)$ of power series type. We shall call the difference $f - g$ a minimal function. A function is then minimal if and only if its corresponding function is 0. The uniqueness statement of the preceding theorem states that if $f(z)$ is minimal, $f + g$ is never minimal if g is of power series type, unless $g = 0$ almost everywhere on γ. To solve the minimum problem of the paper we need only give a complete description of minimal functions. Such a description will be given in Theorem 5, but it will be necessary to prove a preliminary lemma first.

LEMMA 4. *Let $\{f_n(z)\}$ be a sequence of measurable functions on γ, and suppose that $|f_n(z)| \leqq u(z)$ on γ, $n \geqq 1$, where $u(z)$ is integrable on γ. Suppose that $f_n \to f$ almost everywhere on γ. Then if $f_n - g_n$, $f - g$ are minimal functions,*

$$\lim_{n\to\infty} \int_1^z g_n(z) \, dz = \int_1^z g(z) \, dz, \tag{4.1}$$

$$\lim_{n\to\infty} \int_1^z |f_n(z) - g_n(z)| \, |dz| = \int_1^z |f(z) - g(z)| \, |dz|, \tag{4.2}$$

everywhere on γ.

Define F_n, F, G_n, G by

$$(4.3)\qquad \begin{aligned} F_n(z) &= \int_1^z f_n(z)\,dz, \qquad & G_n(z) &= \int_1^z g_n(z)\,dz, \\ F(z) &= \int_1^z f(z)\,dz, \qquad & G(z) &= \int_1^z g(z)\,dz. \end{aligned}$$

The variation of G_n is

$$(4.4)\qquad \begin{aligned} \int_\gamma |g_n|\,|dz| &\leqq \int_\gamma |f_n - g_n|\,|dz| + \int_\gamma |f_n|\,|dz| \\ &\leqq 2\int_\gamma |f_n|\,|dz| \leqq 2\int_\gamma u(z)\,|dz|. \end{aligned}$$

By Helly's theorem, there is then an everywhere convergent subsequence $\{G_{a_n}\}$ of $\{G_n\}: G_{a_n} \to \bar{G}_1(z)$. The function $\bar{G}_1(z)$ is of power series type, and of bounded variation, so it coincides except at possible discontinuity points with an absolutely continuous function $\bar{G}(z)$, also of power series type, with derivative $\bar{g}(z)$. Since $F_{a_n} - G_{a_n} \to F - \bar{G}_1$ on γ,

$$(4.5)\qquad \begin{aligned} \text{Var.}\,(F - \bar{G}) &= \int_\gamma |f - \bar{g}|\,|dz| \leqq \text{Var.}\,(F - \bar{G}_1) \\ &\leqq \liminf_{n\to\infty} \text{Var.}\,(F_{a_n} - G_{a_n}) = \liminf_{n\to\infty} \int_\gamma |f_{a_n} - g_{a_n}|\,|dz| \\ &\leqq \liminf_{n\to\infty} \int_\gamma |f_{a_n} - g|\,|dz| = \int_\gamma |f - g|\,|dz|. \end{aligned}$$

Since $f - g$ is minimal, and since there is essentially only one such g, we must have $\bar{g} = g$ almost everywhere on γ: there must be equality throughout (4.5) and $\bar{G}_1 \equiv \bar{G} \equiv G$. Thus the sequence $\{G_n\}$ has only one possible limiting function, G, so it must converge to G, as was to be shown. Moreover, we have shown that

$$(4.6)\qquad \lim_{n\to\infty} \int_\gamma |f_n - g_n|\,|dz| = \int_\gamma |f - g|\,|dz|.$$

Now if I is any arc of γ not containing 1 in its interior,

$$(4.7)\qquad \underset{I}{\text{Var.}}\,(F - G) \leqq \liminf_{n\to\infty} \underset{I}{\text{Var.}}\,(F_n - G_n),$$

that is,

$$(4.8)\qquad \int_I |f - g|\,|dz| \leqq \liminf_{n\to\infty} \int_I |f_n - g_n|\,|dz|.$$

This is incompatible with (4.6) unless there is always equality in (4.8). Since the equality in (4.8) must hold not only for the original sequence, but also for any subsequence, the "lim inf" can be replaced by "lim": (4.2) is true.

THEOREM 5. *Let $p(z)$ be a function of power series type, of modulus not exceeding* 1 *almost everywhere on γ. Then if $\sigma(z)$ is a non-negative integrable function on γ, which vanishes almost everywhere where $|p(z)| < 1$,*

$$f(z) = \frac{\sigma(z)}{zp(z)} \tag{5.1}$$

is a minimal function, and every minimal function can be obtained in this way.

In other words, *f is minimal if and only if $|f|/zf$ coincides almost everywhere on γ, where $f \neq 0$ with a function of power series type, of modulus not exceeding* 1. We shall prove this in several steps.

(i) Suppose that the minimal function $f(z)$ never vanishes, or vanishes at most on a set of measure 0. Then the fact that, if $n \geqq 0$,

$$\int_\gamma |f(z) + \lambda z^n| \, |dz| \tag{5.2}$$

has its absolute minimum when $\lambda = 0$ implies, according to Theorem 2, that

$$\int_\gamma \frac{|f|}{f} z^n |dz| = 0, \tag{5.3}$$

that is, that $|f|/zf$ is of power series type. Then Theorem 5 is true in this case, in which $|p(z)| = 1$ almost everywhere on γ.

(ii) Let ϵ be any number. Then there is a function $g_\epsilon(z)$, of power series type, such that

$$\varphi_\epsilon(z) = f(z) + \frac{\epsilon}{z} - g_\epsilon(z) \tag{5.4}$$

is minimal. We shall prove that $\varphi_\epsilon(z) = 0$ on at most a set of measure 0 on γ if ϵ is not in some exceptional set which is at most denumerably infinite. Let E_ϵ be the z-set where $\varphi_\epsilon(z) = 0$. Then

$$\frac{\epsilon_1 - \epsilon_2}{z} = g_{\epsilon_1}(z) - g_{\epsilon_2}(z) \qquad (z \in E_{\epsilon_1} \cdot E_{\epsilon_2}), \tag{5.5}$$

or

$$0 = (\epsilon_1 - \epsilon_2) - z[g_{\epsilon_1}(z) - g_{\epsilon_2}(z)] \qquad (z \in E_{\epsilon_1} \cdot E_{\epsilon_2}). \tag{5.6}$$

If $m(E_{\epsilon_1} \cdot E_{\epsilon_2}) \neq 0$, the function on the right in (5.6) is a function of power series type, vanishing on a set of positive measure. Then it vanishes almost everywhere on γ, so (5.5) is true almost everywhere on γ. But if $\epsilon_1 \neq \epsilon_2$ the function on the right in (5.5) is of power series type, whereas that on the left is not, so equality almost everywhere on γ is impossible. Thus if $\epsilon_1 \neq \epsilon_2$, $m(E_{\epsilon_1} \cdot E_{\epsilon_2}) = 0$. Then at most denumerably many sets E_ϵ are of positive measure.

(iii) Let $\{\epsilon_n\}$ be a sequence of constants converging to 0, chosen so that

$\varphi_{\epsilon_n}(z) = 0$ at most on a set of measure 0. Then according to (i), there is a function $p_n(z)$, of power series type, such that

$$|\varphi_{\epsilon_n}(z)| = z\varphi_{\epsilon_n}(z)p_n(z) \tag{5.7}$$

almost everywhere on γ. We can suppose the sequence $\{\epsilon_n\}$ to have been chosen so that the sequence $\{P_n(z)\}$: $\left\{\int_1^z p_n(z)\,dz\right\}$ (a uniformly bounded sequence of functions of uniformly bounded variation) converges everywhere on γ. The limit function $\bar{P}(z)$ is of power series type. Since it is of bounded variation, it coincides on γ except for possible jumps with an absolutely continuous function $P(z)$. We denote by $P(z)$, $P_n(z)$, $p(z)$, $p_n(z)$ the functions analytic in γ, with boundary functions $P(e^{i\theta})$, $P_n(e^{i\theta})$, $p(e^{i\theta}) = P'(e^{i\theta})$, $p_n(e^{i\theta})$ respectively. If we use the Cauchy integral formula, $P_n(z) \to P(z)$ uniformly in every closed subregion of γ. Evidently $p_n(z) = P_n'(z)$, $p(z) = P'(z)$ in γ, and $p_n(z) \to p(z)$ uniformly in every closed subregion of γ. Since $|p_n(z)| \leqq 1$ in γ, $|p(z)| \leqq 1$ in γ also, so $|p(z)| \leqq 1$ almost everywhere on γ.

(iv) To prove (5.1), with the $p(z)$ obtained in (iii), and with $\sigma = |f|$, we shall prove that

$$\int_\gamma |f(z)|\,z^k\,|dz| = \int_\gamma z^{k+1}f(z)p(z)\,|dz|, \tag{5.8}$$

for every integer k. To prove (5.8) it will be sufficient, because of (5.7), to prove

$$\lim_{n\to\infty}\int_\gamma |\varphi_{\epsilon_n}(z)|\,z^k\,|dz| = \int_\gamma |f(z)|\,z^k\,|dz| \tag{5.9}$$

and

$$\lim_{n\to\infty}\int_\gamma z^{k+1}\varphi_{\epsilon_n}(z)p_n(z)\,|dz| = \int_\gamma z^{k+1}f(z)p(z)\,|dz|. \tag{5.10}$$

Now, according to Theorem 4,

$$\lim_{n\to\infty}\int_1^z |\varphi_{\epsilon_n}(z)|\,|dz| = \int_1^z |f(z)|\,|dz| \tag{5.11}$$

on γ, and (5.11) implies (5.9) (integrate (5.9) by parts). To prove (5.10) it is sufficient to prove

$$\lim_{n\to\infty}\int_\gamma z^{k+1}f(z)p_n(z)\,|dz| = \int_\gamma z^{k+1}f(z)p(z)\,|dz| \tag{5.12}$$

and

$$\lim_{n\to\infty}\int_\gamma z^{k+1}g_{\epsilon_n}(z)p_n(z)\,|dz| = 0. \tag{5.13}$$

Now (5.12) is true if $f(z)$ has a continuous derivative (integrate by parts, using the fact that $P_n(z) \to P(z)$), so (5.12) is true for any integrable $f(z)$ since such

an $f(z)$ can be approximated arbitrarily closely (in the sense of distance in L_1) by functions with continuous derivatives. According to Lemma 4,

$$\lim_{n\to\infty} \int_1^z g_{\epsilon_n}(z)\,dz = 0 \tag{5.14}$$

on γ. It follows that

$$\lim_{n\to\infty} \frac{1}{2\pi i} \int_\gamma \frac{g_{\epsilon_n}(t)}{t - z}\,dt = 0 \tag{5.15}$$

uniformly in every closed subregion of $|z| < 1$ (integrate (5.15) by parts). Thus if $g_{\epsilon_n}(z)$ (for $|z| < 1$) denotes the analytic function in (5.15), which has the boundary function $g_{\epsilon_n}(e^{i\theta})$, $g_{\epsilon_n}(z)$, and therefore $g_{\epsilon_n}(z)p_n(z)$, converge to 0 uniformly in every closed subregion of $|z| < 1$. But this means that the derivatives of $g_{\epsilon_n}(z)p_n(z)$ converge to 0 also, and this convergence at $z = 0$ is precisely the content of (5.13) for $k < 0$. The integrals all vanish when $k \geqq 0$, since $g_{\epsilon_n}(z)p_n(z)$ is of power series type. The proof of (5.1) is now complete.

(v) Conversely suppose that $f(z)$ is given by (5.1), with σ, p as described. Let $g(z)$ be any function of power series type. Then $p(z)g(z)$ is also of power series type, since p is bounded, and

$$\left| \int_{f\neq 0} \frac{|f|}{f}\, g\, |dz| \right| = \left| \frac{1}{i} \int_{f\neq 0} p(z)g(z)\,dz \right| \tag{5.16}$$

$$= \left| \int_{f=0} p(z)g(z)\,dz \right| \leqq \int_{f=0} |g(z)|\,|dz|,$$

so that

$$\int_\gamma |f(z) + \lambda g(z)|\,|dz|, \tag{5.17}$$

according to Theorem 2, takes on its minimum value at $\lambda = 0$. Then $f(z)$ is a minimal function, as was to be proved.

The following theorem relates a given f to the g of power series type which makes $f - g$ minimal.

THEOREM 6. *Let $f(z)$ have the form*

$$f(z) = \frac{u(z)}{\prod_1^n (z - \alpha_j)}, \qquad |\alpha_j| < 1, \tag{6.1}$$

where $u(z)$ is of power series type and $n \geqq 1$. Then the function $g(z)$ which makes $f - g$ minimal has the form

$$g(z) = f(z) - \frac{G(z)}{\prod_j (z - \alpha_j)(1 - \bar{\alpha}_j z)}, \tag{6.2}$$

where $G(z)$ is a polynomial of degree not exceeding $2n - 2$, uniquely determined by f. In particular, if $f(z)$ is a rational function with no poles on γ, the corresponding g is also rational.

If $f - g$ is minimal we can write $|f - g| = (f - g)zp$, where $p(z)$ is of power series type. Then $\overline{f - g} = (f - g)z^2p^2$. In this case the latter equation becomes

$$\overline{u - g\prod_j (z - \alpha_j)} = \frac{[u - g\prod_j (z - \alpha_j)]\prod_j (\bar{z} - \bar{\alpha}_j)}{\prod_j (z - \alpha_j)} z^2p^2. \tag{6.3}$$

If G is defined by

$$G(z) = \prod_j (1 - \bar{\alpha}_j z)[u - g\prod_j (z - \alpha_j)], \tag{6.4}$$

then G is of power series type, and, using (6.3), we have

$$z^{2n-2}\overline{G(z)} = G(z)p(z)^2. \tag{6.5}$$

Thus $z^{2n-2}\bar{G}$ is also of power series type so (Lemma 1) G is equal almost everywhere to a polynomial of degree not exceeding $2n - 2$. Since g is uniquely determined, neglecting sets of measure 0, $G(z)$ is a uniquely determined polynomial. We have now verified (6.2). The minimal $f - g$ in this case is a rational function. The next theorem gives some additional information on the character of this rational function, that is, on the character of $G(z)$.

COROLLARY. *If $f(z)$, given by* (6.1), *is minimal, then f is rational:*

$$f(z) = \frac{G(z)}{\prod_j (z - \alpha_j)(1 - \bar{\alpha}_j z)}. \tag{6.6}$$

In fact (6.6) is merely the condition that g vanish identically. More generally, (6.2) shows that if the minimizing g is to be rational, the original f must also be rational.

THEOREM 7. *Let $f(z)$ ($\not\equiv 0$) be a function defined and analytic in the annulus $A: 1/K < |z| < K$ ($K > 1$):*

$$f(z) = \sum_{-\infty}^{\infty} a_n z^n. \tag{7.1}$$

Define $f_1(z)$ by

$$f_1(z) = \sum_{-\infty}^{\infty} \bar{a}_n z^{-n}. \tag{7.2}$$

Then if $f(z)$ is minimal, f_1/f can be continued analytically over the whole plane, and is a rational function, having no poles within γ, with a double zero at the origin. The zeros of $f(z)$ in A are distributed as follows: if $f(\alpha) = 0$ ν times, $1/K < |\alpha| < 1$, then $f(1/\bar{\alpha}) = 0$ $\nu' \geqq \nu$ times, where $\nu' - \nu$ is even. If $f(\alpha) \neq 0$, $1/K < |\alpha| < 1$,

but $f(1/\bar{\alpha}) = 0$ ν times, then ν is even. The zeros on γ, if any, are of even multiplicity.

If $a_n = 0$ for $n > N$, then $f(z)$ is a rational function of z. If $a_n = 0$ for $n < -N$ $(N > 0)$, then $z^N f(z)$ is a polynomial in z of degree not exceeding $2N - 2$.

By definition, $f_1 = \bar{f}$ on γ. According to Theorem 5, $|f|/zf = p$ is of power series type on γ. Then $f_1 = z^2p^2f$ on γ, and p^2, defined in the given annulus by the equation $p^2 = f_1/z^2f$, is meromorphic in the annulus. Moreover, the zeros of f_1 and f cancel each other on γ so p^2 is actually analytic on γ and does not vanish there. Then $p(z)$, defined within γ as the analytic function with boundary function $p(e^{i\theta})$, can be continued analytically across γ. Since $|p(e^{i\theta})| \equiv 1$, $p(z)$ can be continued analytically over the whole extended plane, by the use of the Schwarz reflection principle, and is analytic except for poles. The function f_1/f is therefore a rational function. If $\alpha_1, \cdots, \alpha_n$ are the zeros of $p(z)$ in γ,

$$p(z) = c \prod_j \frac{z - \alpha_j}{\bar{\alpha}_j z - 1}. \tag{7.3}$$

We now have

$$f_1(z) = p(z)^2 z^2 f(z) \tag{7.4}$$

in A. Simple examples (with $p(z) \equiv 1$) show that $f(z)$ may have circles inside and outside γ as natural boundaries. If $f(\alpha) = 0$ $\nu \geqq 0$ times, and if $p(\alpha) = 0$ $\mu \geqq 0$ times, where $1/K < |\alpha| < 1$, $fz^2p^2 = 0$ $\nu + 2\mu$ times. Then $f(1/\bar{\alpha}) = \overline{f_1(\alpha)} = 0$ $\nu + 2\mu$ times. Since $zpf = |f|$ on γ, the zeros of f on γ, if any, must be of even multiplicity. Then the zeros of f are as described in the theorem.

If $a_n = 0$ for $n > N$, $z^N f_1$ is of power series type, and we have, using (7.3) and (7.4),

$$f(z) = \frac{[z^N f_1(z)] \prod_j (\alpha_j z - 1)^2}{z^{N+2} c^2 \prod_j (z - \alpha_j)^2} = \frac{u(z)}{z^{N+2} \prod_j (z - \alpha_j)^2} \tag{7.5}$$

where $u(z)$ is of power series type. Then by the corollary to Theorem 6, $f(z)$ is rational.

If $a_n = 0$ for $n < -N$ $(N > 0)$, $z^N f$ is of power series type, and, using (7.4), $z^{2N-2}\overline{z^N f} = p^2 z^N f$ is also of power series type. Then according to Lemma 1, $z^N f$ coincides almost everywhere on γ with a polynomial of degree not exceeding $2N - 2$. This is the case considered by Riesz, who showed that the distribution of zeros described above for f is sufficient that f be minimal (A is replaced by the whole plane in Riesz' case). Similar considerations, as Riesz remarked, can be carried through for rational functions, to find the distribution of zeros and poles necessary and sufficient that a rational function be minimal.

It would be natural to vary the minimal problem of the paper by supposing that $|f(z)|$, $|p(z)|$ in (0.1) have integrable k-th powers $(k > 1)$, with $p(z)$ still of power series type, and using the k-th power of the integrand in (0.1). This

hypothesis is easier to treat than the case $k = 1$. In fact in the analogue of Theorem 2, with $k > 1$, there cannot be equality in (2.3), so there can only be a single minimum point λ. Equation (2.2) becomes, if g is bounded,

$$\int_{f\neq 0} g \frac{|f|^k}{f} \, | dz | = 0. \tag{2.2''}$$

The existence of a minimizing g is easily proved in the analogue of Theorem 3, and its uniqueness follows from the simpler version of Theorem 2.[7] From (2.2″) with $g = z^n$, $n \geqq 0$, f is minimal if and only if

$$\frac{|f|^k}{f} = zp,$$

where p is of power series type, and $p^{k'}$ is integrable ($k' = k/(k - 1)$) almost everywhere on γ, if we interpret the left side as 0 if $f = 0$. This means that the minimal function f can vanish only on a set of measure 0 unless $f = 0$ almost everywhere on γ, unlike the situation when $k = 1$. A function f is then minimal if and only if it can be written in the form

$$f(z) = \frac{\sigma(z)}{zp(z)} = \frac{| p(z) |^{k'-2}}{z} \overline{p(z)}, \tag{5.1'}$$

where $\sigma(z) \geqq 0$ is integrable, vanishing almost nowhere, and $p(z)$ is of power series type such that $| p | = \sigma^{1/k'}$. The case $k = 2$ is particularly simple, because in this case a function f is minimal if and only if its Fourier expansion contains only negative powers of $1/z$. This can be checked by the direct evaluation of (0.1).

Another natural problem is to minimize (0.1) where $p(z)$ is now to range through the polynomials of a given degree n. The existence of a minimizing polynomial can be proved as in Theorem 3, with some simplifications made possible by the fact that g is a polynomial. The minimizing g is, however, not uniquely determined. In fact, let $n = 0$, $f(e^{i\theta}) = 1$ for $0 \leqq \theta < \pi$, $f = -1$ for $\pi \leqq \theta < 2\pi$. Then $g = c$ minimizes (0.1) for any real constant c between -1 and 1.

University of Illinois.

[7] Cf. a similar proof in J. L. Walsh, *Interpolation and Approximation*, American Mathematical Society Colloquium Publications, vol. 20, New York, 1935, pp. 363–366.

CHANGE OF VELOCITIES IN A CONTINUOUS ERGODIC FLOW

By Warren Ambrose

Introduction. The fact that the proofs of the ergodic theorem and related theorems hold for general measure-preserving transformations has led to the formulation of the theory of flows in terms of measure theory without reference to the classical dynamical systems from which the theory arose.[1] In this paper we consider flows from this standpoint and prove some theorems about continuous ergodic flows. Our main theorem (Theorem 5) asserts that by a change of velocities any such flow can be put into a certain geometrical form; a flow in this form we call a flow built on a measure-preserving transformation. Associated with any flow is a 1-parameter group of unitary operators and the spectral resolution of such groups of operators has proved a useful tool in discussing flows.[2] Because a flow built on a measure-preserving transformation cannot have a purely continuous spectrum, this theorem can be considered as a theorem about what can be done to the spectral resolution by a change of velocities in a flow.

Roughly, our proof proceeds as follows: we show the existence of a set R (called a regular set) with the properties: (1) almost all points go in and out of R infinitely often, and (2) whenever a point gets into either R or CR, it stays there for some time interval of positive length.[3] Each of the sets R and CR is then a set of finite segments of paths of the flow. We then alter the velocities in such a way that each of these segments is traversed in the same fixed length of time. Except for a number of measurability difficulties it is easy to see that the resulting flow is built on a measure-preserving transformation.

Although we start with a continuous flow, simple examples show that we do not in general end up with a continuous flow after such a change of velocities. However, we show that the flow that we obtain will take Borel sets into Borel sets and will be a measurable flow. Since any essential change in velocities will yield a flow under which the original measure will no longer be invariant, it is necessary to find a new measure invariant under the new flow and "properly" related to the original measure. We show that there is a measure equivalent to the original measure (two measures are equivalent if they vanish for the same sets) which is invariant under the new flow.

Received October 14, 1940.

[1] See, for example, [3] and [6]. (Numbers in brackets refer to the bibliography at the end of this paper.)

[2] The idea of using this spectral resolution for studying flows and measure-preserving transformations is due to Koopman [4].

[3] The referee of this paper has pointed out that we have used our hypothesis that the flow be ergodic only in proving the existence of regular sets and that for this purpose it is sufficient to assume the weaker hypothesis that almost all points of the space lie on transitive (dense) paths.

In the first section we state our notation, definitions and assumptions. In the second section we prove the existence of regular sets and use them to define our change of velocities. The last section contains a proof that the altered flow is built on a measure-preserving transformation, that it has the desired measure properties, and that the measure invariant under it is equivalent to the original measure. We conclude the last section with a theorem on the representation of continuous ergodic flows.

1. **Notation, assumptions and definitions.** Throughout this paper Ω will denote a metric separable space on which a measure $m(M)$ is defined having the following properties:

(a) all Borel sets are measurable;

(b) if M is any measurable set, then $m(M) = \inf m(O)$, for O open and including M;

(c) $m(\Omega) < \infty$;

(d) spheres have positive measure;

(e) points have measure zero.

Definition 1. A *measure-preserving transformation* is a 1:1 transformation of Ω into itself with the property that if M is any measurable set, then the sets TM and $T^{-1}M$ are also measurable and $m(M) = m(TM) = m(T^{-1}M)$.

Definition 2. A *flow* is a 1-parameter family T_t $(-\infty < t < \infty)$ of measure-preserving transformations of Ω into itself with the group property: $T_tT_s = T_{t+s}$, for all real t and s.

Definition 3. The flow T_t is *continuous* if the function T_tP is continuous in the two variables t and P. It is a *measurable* flow if for any open set O the (t, P)-set

$$\underset{(t,P)}{[T_tP \in O]}$$

is measurable in the product space of the real line with Ω.

Definition 4. The flow T_t is *ergodic*[4] if the only measurable invariant sets are of measure 0 or are complements of sets of measure 0. (By an invariant set we mean a set M with the property that if $P \in M$ then $T_tP \in M$ for all t.)

Definition 5. Let T_t be a flow on Ω_1 and S_t a flow on Ω_2. T_t is *isomorphic*[5] to S_t if there exists a 1:1 measure-preserving correspondence between Ω_1 and Ω_2 with the property that if P (in Ω_1) corresponds to Q (in Ω_2), then T_tP corresponds to S_tQ for all t.

Definition 6. Let T be a measure-preserving transformation of a space Ω' into itself. Consider the product space $\Omega' \times I$ of Ω' with the interval I: $0 \leqq x < 1$ (taking measure on $\Omega' \times I$ to be the measure defined multiplicatively in terms of the given measure on Ω' and Lebesgue measure on I). Define the flow T_t on $\Omega' \times I$ by

[4] We prefer the word "ergodic" to the more common term "metrically transitive" because it is shorter and because this concept is not a metric concept.

[5] This is a narrower definition of isomorphism than is sometimes used. See [6], for example.

$$T_t(P, x) = (T^{[t+x]}P, t + x - [t + x]) \qquad \text{for } t \geqq 0,$$
$$T_t(P, x) = T_{-t}^{-1}(P, x) \qquad \text{for } t < 0$$

(where $[t + x]$ = integral part of $t + x$). We call T_t the *flow built on* T.

DEFINITION 7. If T_t and S_t are flows on Ω with the property that for each P the paths T_tP and S_tP are the same, both in the sense of being the same point sets and of having these points in the same order, then we say that T_t and S_t are *obtained from each other by a change of velocities.*

We remark that if T_t is a flow built on a measure-preserving transformation, then it has infinitely many eigenfunctions, for the function $F(P, x) = e^{2\pi i n x}$ (n any integer) is an eigenfunction of eigenvalue n. This remark shows that not every flow is built on a measure-preserving transformation and that (for continuous ergodic flows) Theorem 5 implies that by a change of velocities it is possible to obtain a flow with infinitely many numbers in its point spectrum.

Throughout the rest of this paper T_t will denote a fixed continuous ergodic flow on Ω. We shall denote T_tP by P_t. We now define two functions, associated with any set M, which will be fundamental in all our considerations.

If M is any set in Ω, then $f_M(P)$ and $g_M(P)$ are defined by

$$f_M(P) = \begin{cases} \sup \delta & \text{if } P \in M, \\ 0 & \text{if } P \in CM, \end{cases}$$

$$g_M(P) = \begin{cases} \sup (b - a) & \text{if } P \in M, \\ 0 & \text{if } P \in CM, \end{cases}$$

where the first sup is to be taken over those δ's such that $P_{-t} \in M$ for $0 < t < \delta$ and the second sup is to be taken over those a and b for which $a \leqq 0 \leqq b$ and $P_{-t} \in M$ for $a < t < b$. (These functions may be infinite valued.) For $P \in M$ the function $f_M(P)$ represents the length of time that a particle now at P has been in M since its last entry into M, while $g_M(P)$ represents the length of time the particle has been in M since its last entry plus the length of time it will remain in M before leaving.

2. Regular sets and change of velocities.

In this section we begin by proving some lemmas about the functions $f_M(P)$, $g_M(P)$ just defined. Then we define the term regular set and prove that every open set contains an open regular set. Next we define the change of velocities along the paths of T_t to obtain the 1-parameter group S_t which we are after. We conclude this section with a proof that S_t takes Borel sets into Borel sets.

LEMMA 1. *If O is open, then $f_O(P)$ and $g_O(P)$ are l.s.c.*

Proof. It is sufficient to prove that $f_O(P)$ is l.s.c. because if we denote by $f_O'(P)$ the corresponding function for the flow T_{-t}, then $f_O'(P)$ is l.s.c. and hence $g_O(P) = f_O(P) + f_O'(P)$ is l.s.c.

Since $f_O(P)$ is obviously l.s.c. at points of CO, it is sufficient to prove that if $P^n \epsilon O$, $P \epsilon O$, $P^n \to P$, $f_O(P^n) \to k$ $(k \leqq \infty)$, then $f_O(P) \leqq k$. Let $a_n = f_O(P^n)$. Then $P^n_{-a_n} \epsilon' O$ and $P^n_{-a_n} \to P_{-k} \epsilon' O$. Since $k \geqq 0$, we have $f_O(P) \leqq k$.

LEMMA 2. *If F is closed, then $f_F(P)$ and $g_F(P)$ are Borel functions.*[6]

Proof. Choose open sets O_n decreasing to F as $n \to \infty$. Then it is easily seen that $f_{O_n}(P) \to f_F(P)$, $g_{O_n}(P) \to g(P)$, as $n \to \infty$.

DEFINITION 8. A measurable set R is *regular* if it has the following properties:

(a) Ω can be subdivided into disjunct invariant Borel sets Ω_1 and Ω_2 such that $m(\Omega_2) = 0$ and the path of every point of Ω_1 passes through R;

(b) for each $P \epsilon R$ the t-set for which $P_t \epsilon R$ is a sum of finite open intervals whose complement is a sum of finite non-degenerate closed intervals.

DEFINITION 9. A measurable set M is semi-regular if for each $P \epsilon M$ the t-set for which $P_t \epsilon M$ is open.

Let M be any semi-regular set and define the sets M^+ and M^- by

$$M^+ = M + \underset{P}{[}\text{there exist } t_n \downarrow 0 \text{ such that } P_{-t_n} \epsilon M] = \overline{\lim_{t \downarrow 0}} M_t,$$

$$M^- = M + \underset{P}{[}\text{there exist } t_n \downarrow 0 \text{ such that } P_{t_n} \epsilon M] = \overline{\lim_{t \uparrow 0}} M_t.$$

LEMMA 3. *If M is semi-regular and if $f_M(P)$ and $g_M(P)$ are Borel functions, then $f_{M^+}(P)$, $f_{M^-}(P)$, $g_{M^+}(P)$, $g_{M^-}(P)$ are Borel functions.*

Proof. This theorem is a consequence of the following easily verified relations:

$$\begin{aligned} f_{M^+}(P) &= \lim f_M(P_{-1/n}); & f_{M^-}(P) &= \lim f_M(P_{1/n}) \\ g_{M^+}(P) &= \lim g_M(P_{-1/n}); & g_{M^-}(P) &= \lim g_M(P_{1/n}) \end{aligned} \qquad (u \to \infty).$$

THEOREM 1. *Every (non-empty) open set contains a (non-empty) open regular set.*

Proof. Let O be the open set. Choose an open set O_1 included in O and such that $O - O_1$ contains an open set O_2. Choose $\delta > 0$ such that

$$\underset{P}{[}g_{O_1}(P) > \delta]$$

is non-empty. Because $g_{O_1}(P)$ is l.s.c., this set is open and hence

$$O' = \underset{P}{[}g_{O_1}(P) > \delta]$$

is open and non-empty.

Now define R by

$$R = O' \cdot \underset{P}{[}f_{O_1}(P) > \tfrac{1}{2}\delta].$$

Then R is open because $f_{O_1}(P)$ is l.s.c.; we shall now show that R is regular.

[6] By a Borel function we mean a function $h(P)$ with the property that for every real number k the P-set $[h(P) > k]$ is a Borel set.

Define Ω_1 and Ω_2 by

$$\Omega_1 = \left(\prod_{n=1}^{\infty} \sum_{t \geqq n} T_t O'\right)\left(\prod_{n=1}^{\infty} \sum_{t \leqq -n} T_t O'\right)\left(\prod_{n=1}^{\infty} \sum_{t \geqq n} T_t O_2\right)\left(\prod_{n=1}^{\infty} \sum_{t \leqq -n} T_t O_2\right),$$

$$\Omega_2 = C\Omega_1$$

(Ω_2 may be empty) so that Ω_1 consists of those points whose paths enter both O' and O_2 for infinitely large t and for negatively infinitely large t. Ω_1 and Ω_2 are obviously invariant Borel sets whose sum is Ω. Because our flow is ergodic, each of the four sets that were multiplied to obtain Ω_1 contains almost all of Ω; hence Ω_1 contains almost all of Ω, and $m(\Omega_2) = 0$. If $P \epsilon \Omega_1$, then the path of P passes through O' and hence through R. From the definition of Ω_1 we see that paths of points in Ω_1 go in and out of O', and hence of R, for infinitely large, and negatively infinitely large t; from the definitions of O' and R we see that if $P \epsilon R$ or $P \epsilon CR$, then there is a t-interval (about O) of length at least $\frac{1}{2}\delta$ for which P_t lies in the same set. Hence R is regular.

COROLLARY. *The regular set R of Theorem* 1 *may be chosen with the following property: there exists an $\eta > 0$ such that if $P \epsilon R\Omega_1$ (or $P \epsilon CR \cdot \Omega_1$) then for some t-interval containing* 0 *and of length at least η the points P_t will also lie in $R \cdot \Omega_1$ (or in $CR \cdot \Omega_1$).*

We are now ready to change the velocities along the paths of T_t. We first choose a fixed open regular set R and in terms of it define the new 1-parameter group S_t. *In doing this we shall consider only the set Ω_1 associated with R*; on this space T_t is still a continuous ergodic flow, so our previous theorems hold for it. To define S_t we shall first define a transformation Q and then define S_t to be $Q^{-1}T_tQ$. (It is easy to see that every change of velocities can be defined in this way.) The transformation Q will take each path into itself in a 1:1 order-preserving way (the flow T_t orders the points on each path; it is this order that is preserved). This definition shows immediately that the S_t form a group, i.e., that $S_{t+s} = S_tS_s$; also each S_t is a 1:1 transformation of Ω_1 into itself because each of Q, T_t and Q^{-1} has this property. We remark that in general the transformation Q will not preserve the measurability of sets in Ω_1 though we shall show that each S_t does. In showing that S_t is built on a measure-preserving transformation we shall have to "factor" our space Ω_1 into a product of the real line with some subspace Ω' of Ω_1; i.e., we shall have to find a subspace Ω' whose product space with (0, 1) is—with respect to measure properties—the same as Ω_1. We shall take Ω' to be the set of points at which the paths of our flow enter the regular set R; this subspace serves this purpose because it is a cross section of the flow with the property that under S_t the length of time between successive crossings of Ω' will be 1 for every point of Ω_1.

Now let R be the fixed open regular set. We shall define Q separately along each path Σ. Pick a point P^* on Σ; then each point of Σ is of the form P_t^*. Let the t-intervals for which $P_t^* \epsilon R$ be denoted by (a_n, b_n); then map the real line into itself in an order-preserving way so that (a_n, b_n) is mapped linearly

on $(n, n + \frac{1}{2})$ while (b_n, a_{n+1}) is mapped linearly on $(n + \frac{1}{2}, n + 1)$. Denote this mapping by $F(t)$. Then define Q for $P \in \Sigma$ by $QP_t^* = P_{F(t)}^*$. Doing this for each path, we have Q defined for all points of Ω_1. Now define S_t (for any $P \in \Omega_1$) by $S_tP = Q^{-1}T_tQP$. We note that this definition of Q depended upon which point P^* of Σ we chose in order to represent every point of Σ as a P_t^* and also upon the labelling of the intervals (a_n, b_n). Theorem 2 below shows, however, that the definition of S_t is independent of these choices.

We define Ω' to be $\Omega_1 \cdot (R^- - R)$ so that Ω' consists of these points at which the paths of our flow enter R. We shall frequently denote points of Ω' by P'. Now we define the function $j(P)$ by

$$j(P) = \begin{cases} \dfrac{1}{2}\dfrac{f_{R^-}(P)}{g_{R^-}(P)} & \text{for } P \in R^-, \\[2ex] \dfrac{1}{2} + \dfrac{1}{2}\dfrac{f_{CR^-}(P)}{g_{CR^-}(P)} & \text{for } P \in CR^-. \end{cases}$$

The function $j(P)$ will be very important throughout the rest of this paper; the next theorem shows that $j(P)$ measures how far a particle now at P is through R and CR since its last entry into R, where the measurement is made in terms of S_t. The fact (an immediate consequence of its definition) that $j(P)$ is a Borel function then practically ensures that S_t will have all the measurability properties that we desire of it.

THEOREM 2. *If $P = S_\beta P'$, where $P' \in \Omega'$ and $0 \leqq \beta < 1$, and if $S_\alpha P' \notin \Omega'$ for all α satisfying $0 < \alpha < \beta$, then $j(P) = \beta$.*

Proof. We shall prove this for $0 \leqq \beta \leqq \frac{1}{2}$; the proof for $\frac{1}{2} < \beta < 1$ is essentially the same. Let P'' be the first point beyond P' on the path of P' which lies in CR. Then, writing $a(P)$ for $f_{R^-}(P)$ and $b(P)$ for $g_{R^-}(P)$, we have

$$P = T_{a(P)}P', \qquad P'' = T_{b(P)}P'.$$

Let ${}^1P = QP'$, ${}^2P = QP$ and ${}^3P = QP''$ (where Q is the transformation in terms of which S_t was defined as $Q^{-1}T_tQ$). From $P = S_\beta P' = Q^{-1}T_\beta QP'$ we see that ${}^2P = T_\beta\, {}^1P$. Because, for $0 \leqq t \leqq b(P)$, we have $QP_t' = {}^1P_{G(t)}$, where $G(t)$ is the linear mapping of $(0, b(P))$ into $(0, \frac{1}{2})$, it follows that

$$\beta = \frac{1}{2}\frac{f_{R^-}(P)}{g_{R^-}(P)} = j(P).$$

LEMMA 4. *If $h(P)$ is a Borel function, then the transformation*

$$P \rightarrow T_{h(P)}P$$

is a Borel transformation.[7]

[7] By a Borel transformation we mean a transformation S with the property that if O is any open set, then the P-set $[SP \in O]$ is a Borel set.

Proof. Suppose first that $h(P)$ takes only a finite number of values, the values $a_1, \cdots, a_n$ on the sets $M_1, \cdots, M_n$. Then, if N is any Borel set,

$$\underset{P}{[}T_{h(P)}P \in N] = M_1 \cdot T_{-a_1}N + \cdots + M_n \cdot T_{-a_n}N,$$

so the lemma is true for such $h(P)$.

Now let $h(P)$ be any Borel function. Then let $h_n(P)$ be a sequence of Borel functions, each taking only a finite number of values, and such that $h_n(P) \to h(P)$ $(n \to \infty)$. Then each $T_{h_n(P)}P$ is a Borel transformation, and because T_t is continuous, $T_{h_n(P)}P \to T_{h(P)}P$; hence $T_{h(P)}P$ is a Borel transformation.

THEOREM 3. *Each S_t takes Borel sets into Borel sets.*

Proof. We shall show that if $0 \leqq \delta < \frac{1}{2}$, then S_δ is a Borel transformation, thus showing that for such δ $S_{-\delta}$ takes Borel sets into Borel sets. By the group property of the S_t it will then follow that each S_{-t}, for $t \geqq 0$, takes Borel sets into Borel sets. Similarly, it may be shown that each S_t, for $t < 0$, takes Borel sets into Borel sets.

Let δ be a fixed number, with $0 \leqq \delta < \frac{1}{2}$. Then $S_\delta R^-$ is a Borel set because

$$S_{-\delta}R^- = \underset{P}{[}1 - \delta \leqq j(P) < 1] + \underset{P}{[}0 \leqq j(P) < \tfrac{1}{2} - \delta],$$

so the sets $R^- \cdot S_{-\delta}R^-$, $R^- \cdot CS_{-\delta}R^-$, $CR^- \cdot S_{-\delta}R^-$, $CR^- \cdot CS_{-\delta}R^-$ are Borel sets.

Now write $c(P)$ for $g_{R^-}(P) - f_{R^-}(P)$ and $d(P)$ for $g_{CR^-}(P) - f_{CR^-}(P)$. Then defining $h(P)$ by

$$h(P) = \begin{cases} 2\delta g_{R^-}(P) & \text{for } P \in R^- \cdot S_{-\delta}R^-, \\ g_{R^-}(P) + 2[j(P) + \delta - \frac{1}{2}]g_{CR^-}(T_{c(P)}P) - f_{R^-}(P) & \text{for } P \in R^- \cdot CS_{-\delta}R^-, \\ 2\delta g_{CR^-}(P) & \text{for } P \in CR^- \cdot CS_{-\delta}R^-, \\ g_{CR^-}(P) + 2[j(P) + \delta - \frac{1}{2}]g_{R^-}(T_{d(P)}P) - f_{CR^-}(P) & \text{for } P \in CR^- \cdot S_{-\delta}R^-, \end{cases}$$

it can be shown by a direct computation that $S_\delta P = T_{h(P)}P$. Since $h(P)$ is clearly a Borel function, it follows from Lemma 4 that S_δ is a Borel transformation.

3. **Proof that S_t is built on a measure-preserving transformation.** If $P \in \Omega_1$, then we make correspond to P the coördinate pair (P', x), where $P' \in \Omega'$, $S_x P' = P$ and $j(P) = x$. This establishes a 1:1 correspondence between Ω_1 and the product space of Ω' with the interval I: $0 \leqq x < 1$. Now we shall find a Borel field in Ω' whose product Borel field with the Borel sets in $0 \leqq x < 1$ corresponds to the Borel sets in Ω_1. Then we shall put a measure on Ω' such that S_1 (S_1 obviously takes Ω' into itself) is a measure-preserving transformation on Ω'. It is obvious that the flow on $\Omega' \times I$ which is built on S_1 on Ω' corresponds to S_t and hence that this product measure is invariant under S_t. We shall show also that this product measure is equivalent to the original measure, thus completing the proof that S_t is isomorphic to a flow built on a measure-preserving transformation.

Let $\mathcal{B}$ be the collection of Borel sets in Ω_1, and let $\mathcal{B}'$ be the collection of sets of the form $\Lambda \cdot \Omega'$, where $\Lambda \in \mathcal{B}$. Let $\mathcal{B}_1$ be the sets of $\Omega' \times I$ which are in the product Borel field of $\mathcal{B}'$ and the Borel sets of $0 \leq x < 1$. Let $\mathcal{B}_2$ be the sets of Ω_1 which correspond to sets of $\mathcal{B}_1$ under the above correspondence.

LEMMA 5. $\mathcal{B} = \mathcal{B}_2$.

Proof. If $\Lambda' \subset \Omega'$, we shall call the set $\Lambda = \sum_{0 \leq t < 1} S_t \Lambda'$ the strip based on Λ'. We see that $\Lambda' \in \mathcal{B}'$ if and only if $\Lambda \in \mathcal{B}$, because if $\Lambda \in \mathcal{B}$ then $\Lambda' = \Lambda \cdot \Omega'$, while if $\Lambda' \in \mathcal{B}'$ then Λ is the origin of Λ' under $T_{-h(P)}P$, where

$$h(P) = \begin{cases} f_{R^-}(P) & \text{if } P \in R^-, \\ f_{CR^-}(P) + g_{R^+}(T_{-f_{CR^-}(P)}P) & \text{if } P \in CR^-, \end{cases}$$

so by Lemma 4 $\Lambda \in \mathcal{B}$.

The determining sets of $\mathcal{B}_1$ are sets of the form $\Lambda' \times (a, b)$ where $\Lambda' \in \mathcal{B}'$; the correspondent of such a set in Ω_1 is the set

$$\Lambda \cdot \underset{P}{[}a < j(P) < b],$$

where Λ is the strip based on Λ', so the sets of this form are a determining collection for $\mathcal{B}_2$. Because such sets are in $\mathcal{B}$, it follows that $\mathcal{B}_2 \subset \mathcal{B}$. Also because such sets obviously go, under each S_t, into sets of $\mathcal{B}_2$, it follows that each S_t takes sets of $\mathcal{B}_2$ into sets of $\mathcal{B}_2$.

We now want to show that $\mathcal{B} \subset \mathcal{B}_2$. To prove this it is sufficient to show that if O is any open set in R, then $O \in \mathcal{B}_2$. This is sufficient because assuming it and letting Λ be any set in $\mathcal{B}$ we can write

$$\Lambda = \Lambda R + \Lambda \cdot S_{\frac{1}{2}} R + \Lambda \cdot [R^+ - R] + \Lambda \cdot [R^- - R]$$

and then show these sets to be in $\mathcal{B}_2$ as follows:

(a) ΛR is a Borel set in R; since open sets in R belong to $\mathcal{B}_2$, it follows that Borel sets in R belong to $\mathcal{B}_2$, and hence $\Lambda R \in \mathcal{B}_2$.

(b) $\Lambda \cdot S_{\frac{1}{2}} R$ is a Borel set, so (by Theorem 2) $S_{-\frac{1}{2}}(\Lambda \cdot S_{\frac{1}{2}} R) = S_{-\frac{1}{2}}\Lambda \cdot R$ is a Borel set; since it is in R, it belongs to $\mathcal{B}_2$. Hence (by the last paragraph) $S_{\frac{1}{2}}(S_{-\frac{1}{2}}\Lambda \cdot R) = \Lambda \cdot S_{\frac{1}{2}} R \in \mathcal{B}_2$.

(c) $\Lambda \cdot (R^- - R) = \Lambda \cdot \Omega'$ belongs to $\mathcal{B}'$ and hence to $\mathcal{B}_2$.

(d) $\Lambda \cdot [R^+ - R] = S_{\frac{1}{2}}(S_{-\frac{1}{2}}\Lambda)(S_{-\frac{1}{2}}[R^+ - R]) = S_{\frac{1}{2}}(S_{-\frac{1}{2}}\Lambda)\Omega'$.

Again because the S_t take sets of $\mathcal{B}$ into sets of $\mathcal{B}$ and sets of $\mathcal{B}_2$ into sets of $\mathcal{B}_2$, it follows that this set belongs to $\mathcal{B}_2$.

To conclude the proof of this lemma we shall now let O be an arbitrary open set in R and show that $O \in \mathcal{B}_2$. Let $\{O_m\}$ be a denumerable basis for the open sets in Ω and let $\{O'_m\}$ be the non-empty sets of the form $O_m \cdot \Omega'$; then let Z_m be the strip based on O'_m. Define the sets M_n^k, ${}^m M_n^k$ by

$$M_n^k = \underset{P}{[}k2^{-n} \leq g_{R^-}(P) < (k+1)2^{-n}] \quad (n = 1, 2, \cdots; k = -\infty, \cdots, \infty),$$

$${}^m M_n^k = Z_m \cdot M_n^k \quad (n = 1, 2, \cdots; k = -\infty, \cdots, \infty).$$

Because $M_n^k + S_{\frac{1}{2}}M_n^k$ is obviously a strip and in $\mathfrak{B}$, and because

$$M_n^k = \underset{P}{[}0 \leqq j(P) < \tfrac{1}{2}]\cdot\{M_n^k + S_{\frac{1}{2}}M_n^k\},$$

we see that each M_n^k is also in $\mathfrak{B}_2$. Now consider the following sums, where the summation is extended over all sets of the indicated form which are included in the given open set O:

$$\Lambda_n = \sum_{r,r',m,k} \underset{P}{[}r < j(P) < r']\cdot{}^mM_n^k$$

(here the r, r' range through the rational numbers, k ranges through all integers, and m ranges through the positive integers). We shall show that $\Lambda_n \to O$ as $n \to \infty$. Since $\Lambda_n \subset O$, we must show only that if $P \in O$ then $P \in \Lambda_n$ for sufficiently large n. In what follows P will denote a fixed but arbitrary point of O. We shall prove separately the statements (i), (ii) and (iii) below and then use them to show that $P \in \Lambda_n$ for large n. For simplicity in notation we write $f(P)$ for $f_{R^-}(P)$ and $g(P)$ for $g_{R^-}(P)$ throughout the remainder of this proof.

(i) There exist a neighborhood $N(P)$ and a number $a > 0$ such that if $P^* \in N(P)$ then $f_O(P^*) > a$, $g_O(P^*) - f_O(P^*) > a$. This follows immediately from the fact that these functions are l.s.c. and positive at P.

(ii) If $\epsilon > 0$. then there is a neighborhood of P, $N(P)$, included in O, and an integer K_1 such that if $n \geqq K_1$, $P \in M_n^k$, $P^* \in N(P)\cdot M_n^k$, then $|f(P) - f(P^*)| < \epsilon$. First we choose K_1 such that $2^{-K_1} < \frac{1}{2}\epsilon$. Then

(a) if P, P^* belong to the same M_n^k for $n \geqq K_1$, then $|g(P) - g(P^*)| < \frac{1}{2}\epsilon$.
Next we choose $N'(P)$ such that

(b) $f(P^*) > f(P) - \frac{1}{2}\epsilon$ for $P^* \in N'(P)$.
Lastly we choose $N''(P)$ such that

(c) $g(P^*) - f(P^*) > g(P) - f(P) - \frac{1}{2}\epsilon$ for $P^* \in N''(P)$.
For the $N(P)$ to satisfy statement (ii) we choose any $N(P)$ included in $N'(P)\cdot N''(P)$. Then if $n \geqq K_1$, $P \in M_n^k$ and $P^* \in N(P)\cdot M_n^k$, we have

$$\begin{aligned} f(P^*) &> f(P) - \tfrac{1}{2}\epsilon && \text{by (b)},\\ f(P) - f(P^*) &> g(P) - g(P^*) - \tfrac{1}{2}\epsilon && \text{by (c)},\\ &> -\epsilon && \text{by (a)}. \end{aligned}$$

Hence $|f(P) - f(P^*)| < \epsilon$ as we desired to show.

(iii) If $\epsilon > 0$, then there is a neighborhood of P, $N(P)$, included in O, and an integer K_1 such that if $n \geqq K_1$, $P \in M_n^k$ and $P^* \in N(P)M_n^k$, then $|j(P) - j(P^*)| < \epsilon$. Choose $N'(P)$ included in O and $a > 0$ such that if $P^* \in N'(P)$, then $P_t^* \in O$ for $-a < t < a$. Choose $N''(P) \subset O$ and an integer K_2 such that if $P^* \in N''(P)$, $n \geqq K_2$, and P, P^* belong to the same M_n^k, then $|f(P) - f(P^*)| < \frac{1}{2}\epsilon a$. Then let $N(P)$ be any neighborhood included in $N'(P)\cdot N''(P)$ and let K_1 be any integer not less than K_2 and such that $2^{-K_1} < \frac{1}{2}a\epsilon$. We now show that this choice of $N(P)$ and K_1 satisfies (iii):

$$\begin{aligned} &|f(P)g(P^*) - f(P^*)g(P)|\\ &\qquad \leqq |f(P)g(P^*) - f(P)g(P)| + |f(P)g(P) - f(P^*)g(P)|\\ &\qquad \leqq f(P)\,|g(P^*) - g(P)| + g(P)\,|f(P) - f(P^*)|. \end{aligned}$$

Because of our choice of $N(P)$ and K_1 we see that for $n \geqq K_1$ and $P, P^* \epsilon N(P)M_n^k$ this expression is less than or equal to $g(P) \cdot a\epsilon$. Dividing by $g(P)g(P^*)$ we get

$$\left| \frac{f(P)}{g(P)} - \frac{f(P^*)}{g(P^*)} \right| < \frac{a\epsilon}{g(P^*)} \leqq \epsilon.$$

This proves (iii).

We are now ready to show that $P \epsilon \Lambda_n$ for large n. Choose first an $N'(P) \subset O$, and an $a > 0$, in accordance with (i). Then choose r, r' rational and such that $r < \beta < r'$ (where $\beta = j(P)$), and $r' - r < a/[g(P) + 1]$. Then choose an $N''(P)$ and an integer K_2 in accordance with (iii), with ϵ = minimum of $\beta - r$ and $r' - \beta$. Now let K_1 be an integer not less than K_2 and such that $2^{-K_1} < \frac{1}{2}a$. We shall show that $P \epsilon \Lambda_n$ for $n \geqq K_1$.

Let $N(P)$ be any neighborhood included in $N'(P) \cdot N''(P)$. Now $T_y P' = P$, where $P' \epsilon \Omega'$ and $y = f(P)$. Because T_y is continuous, we can find a neighborhood of P' and hence a set O_m' containing P' such that $T_y O_m'$ is included in $N(P)$. Now let n be any integer not less than K_1; choose k such that $P \epsilon M_n^k$. We shall show that with these choices of r, r', m, k the set

$$\Lambda = \underset{P^*}{[} r < j(P^*) < r'] \cdot {}^m M_n^k$$

is included in O, thus showing that $P \epsilon \Lambda_n$, for $n \geqq K_1$.

We note that

(a) if $P^* \epsilon {}^m M_n^k \cdot N(P)$, then $P^* \epsilon \Lambda$. This is true because if P^* is in $N(P)$ then $| j(P^*) - \beta | < \min (\beta - r, r' - \beta)$ and hence $r < j(P^*) < r'$.

(b) If $P^* \epsilon \Lambda$, then for some t, with $-f(P^*) < t < g(P^*) - f(P^*)$, $T_t P^* \epsilon {}^m M_n^k \cdot N(P)$. To see this write $P^* = T_z P^{*\prime}$, where $P^{*\prime} \epsilon \Omega'$ and $z = f(P^*)$. Then $P^{*\prime} \epsilon O_m'$ (because $P^* \epsilon {}^m M_n^k$) and hence $T_y P^{*\prime} \epsilon N(P)$, i.e., $T_{y-z} P^* \epsilon N(P)$. We want to show that $-f(P^*) < y - z < g(P^*) - f(P^*)$, i.e., that $0 < y < g(P^*)$. Obviously $y > 0$ and since $P \epsilon N(P)$ we have

$$y \leqq g(P) - a < g(P^*) - \tfrac{1}{2}a < g(P^*).$$

(c) If $P^* \epsilon {}^m M_n^k \cdot N(P)$ and if $T_t P^* \epsilon \Lambda$, where $-f(P^*) < t < g(P^*) - f(P^*)$, then $-a < t < a$. This is proved as follows:

$$rg(P^*) < f(P^*) < r'g(P^*),$$

$$rg(P^*) < f(P^*) + t < r'g(P^*),$$

$$-a \leqq \frac{-a}{g(P) + 1} g(P^*) < (r - r')g(P^*)$$

$$< t < (r' - r)g(P^*) < \frac{a}{g(P) + 1} g(P^*) \leqq a.$$

Now let P^* be any point in Λ. Then by (b) and (c) $P^* = T_t \bar{P}^*$, where $\bar{P}^* \epsilon {}^m M_n^k N(P)$ and $-a < t < a$. Since $\bar{P}^* \epsilon N(P)$, it follows that $P^* = T_t \bar{P}^* \epsilon O$, because of the original choice of $N(P)$. Hence $\Lambda \subset O$. Hence $P \epsilon \Lambda_n$, for $n \geqq K_1$.

We now define the measure $\mu(M)$ which will turn out to be the invariant measure for S_t. Letting $m(M)$ be the original measure on Ω we define $\mu(M)$, for any measurable set M, by

$$\mu(M) = \int_M \frac{1}{g_{R^-}(P) + g_{CR^-}(P)}\, dm.$$

Because this integrand is bounded by $2/\delta$ (where δ is the number used in defining the regular set R, i.e., $2/\delta$ is the η of the corollary to Theorem 1) and is non-negative, $\mu(M)$ is a finite measure equivalent to $m(M)$. The purpose of the next three lemmas is to show that $\mu(M)$ is the direct product measure of a measure on Ω' which is invariant under S_1 and Lebesgue measure on $0 \leqq x < 1$.

Let Λ be a Borel set which is a strip; then we shall call the sets

$$\Lambda \cdot \underset{P}{[}0 \leqq j(P) < \tfrac{1}{2}], \qquad \Lambda \cdot \underset{P}{[}\tfrac{1}{2} \leqq j(P) < 1]$$

half-strips. The former we shall call an *upper* half-strip, the latter a *lower* half-strip.

LEMMA 6. *If M is a half-strip, then $\mu(M) = \mu(S_{\frac{1}{2}}M)$.*

Proof. We note first that each set $N = [j(P) = \text{constant}]$ has measure 0 because for $0 \leqq t < \frac{1}{2}\delta$ the sets T_tN are mutually disjunct and are all Borel sets of the same m-measure.

We shall prove this in case M is a lower half-strip; if M is an upper half-strip, the proof is essentially the same. With each point P of Ω we associate the point P^* defined by

$$P^* = P^*(P) = S_{\frac{1}{2}-j(P)}P.$$

Now we define the sets Λ^n_{jk}, ${}_*\Lambda^n_{jk}$, ${}^*\Lambda^n_{jk}$ by

$$\Lambda^n_{jk} = \underset{P}{[}j2^{-n} \leqq g_{R^+}(P^*) < (j+1)2^{-n}] \cdot \underset{P}{[}k2^{-n} \leqq g_{CR^-}(P^*) < (k+1)2^{-n}]$$

$$\cdot \underset{P}{[}g_{R^+}(P) - f_{R^+}(P) < j2^{-n}] \cdot \underset{P}{[}f_{CR^-}(P) < k2^{-n}],$$

$${}_*\Lambda^n_{jk} = M \cdot \Lambda^n_{jk}, \qquad {}^*\Lambda^n_{jk} = S_{\frac{1}{2}}M \cdot \Lambda^n_{jk}$$

and we define the functions ${}^*g^n(P)$, ${}_*g^n(P)$ by

$${}_*g^n(P) = \begin{cases} j2^{-n} & \text{if } P \in \sum_k {}_*\Lambda^n_{jk}, \\ 0 & \text{if } P \in' \sum_k {}_*\Lambda^n_{jk}, \end{cases}$$

$${}^*g^n(P) = \begin{cases} k2^{-n} & \text{if } P \in \sum_j {}^*\Lambda^n_{jk}, \\ 0 & \text{if } P \in' \sum_j {}^*\Lambda^n_{jk}. \end{cases}$$

Then it is obvious that for $P \in \Lambda = M + S_{\frac{1}{2}}M$

$${}^*g^n(P) \to g_{CR^-}(P), \qquad {}_*g^n(P) \to g_{R^+}(P) \qquad (n \to \infty).$$

Now define the functions $\varphi_n(P)$ by

$$\varphi_n(P) = \begin{cases} \dfrac{1}{{}^*g^n(P) + {}_*g^n(P)} & \text{for } P \text{ such that } {}^*g^n(P) + {}_*g^n(P) > 0, \\ 0 & \text{for other } P. \end{cases}$$

Then for $P \in \Lambda - N$, where N is the null set $\underset{P}{[}j(P) = \frac{1}{2}] + \underset{P}{[}j(P) = 0]$, we have

$$\varphi_n(P) \to \frac{1}{g_{R^-}(P) + g_{CR^-}(P)} \qquad (n \to \infty).$$

Because the $\varphi_n(P)$ are uniformly bounded for large n (more precisely, if K is such that $2^{-K} < \frac{1}{2}\delta$, then $\varphi_n(P) \leqq 4/\delta$ for $n \geqq K$), we can conclude that

$$\int_M \varphi_n(P)\,dm \to \int_M \frac{1}{g_{R^-}(P) + g_{CR^-}(P)}\,dm,$$

$$\int_{S_{\frac{1}{2}}M} \varphi_n(P)\,dm \to \int_{S_{\frac{1}{2}}M} \frac{1}{g_{R^-}(P) + g_{CR^-}(P)}\,dm.$$

Hence it will be sufficient to show that, for each n,

$$(a) \qquad \int_M \varphi_n(P)\,dm = \int_{S_{\frac{1}{2}}M} \varphi_n(P)\,dm.$$

Because of

$$\int_M \varphi_n(P)\,dm = \sum_{j,k} \int_{{}_*\Lambda^n_{jk}} \frac{1}{{}_*g^n(P)}\,dm,$$

$$\int_{S_{\frac{1}{2}}M} \varphi_n(P)\,dm = \sum_{j,k} \int_{{}^*\Lambda^n_{jk}} \frac{1}{{}^*g^n(P)}\,dm,$$

it is sufficient to show that for fixed j, k:

$$(\beta) \qquad \int_{{}_*\Lambda^n_{jk}} \frac{1}{{}_*g^n(P)}\,dm = \int_{{}^*\Lambda^n_{jk}} \frac{1}{{}^*g^n(P)}\,dm.$$

Now

$$\int_{{}_*\Lambda^n_{jk}} \frac{1}{{}_*g^n(P)}\,dm = \int_{{}_*\Lambda^n_{jk}} \frac{2^n}{j}\,dm = \frac{2^n}{j}\,m({}_*\Lambda^n_{jk}),$$

$$\int_{{}^*\Lambda^n_{jk}} \frac{1}{{}^*g^n(P)}\,dm = \int_{{}^*\Lambda^n_{jk}} \frac{2^n}{k}\,dm = \frac{2^n}{k}\,m({}^*\Lambda^n_{jk}).$$

Hence to prove (β) it will be sufficient to show that

$$(\gamma) \qquad m({}_*\Lambda^n_{jk}) = \frac{k}{j}\,m({}^*\Lambda^n_{jk}).$$

To show this we subdivide ${}_*\Lambda^n_{jk}$ into the sets

$${}_*\Lambda(\alpha) = {}_*\Lambda^n_{jk}\cdot\underset{P}{[}\alpha 2^{-n} \leqq g_R(P) - f_R(P) < (\alpha + 1)2^{-n}] \qquad (\alpha = 0, \cdots, j - 1)$$

and subdivide ${}^*\Lambda^n_{jk}$ into the sets

$${}^*\Lambda(\beta) = {}^*\Lambda^n_{jk}\cdot\underset{P}{[}\beta 2^{-n} \leqq f_{CR}(P) < (\beta + 1)2^{-n}] \qquad (\beta = 0, \cdots, k - 1).$$

Then $m({}_*\Lambda(\alpha)) = m({}^*\Lambda(\beta))$ for each α, β since these sets are all transforms of one another under members of T_t (more specifically, ${}^*\Lambda(\beta) = T_{(j-\alpha+\beta)/2^n}\,{}_*\Lambda(\alpha))$. Since the number of ${}_*\Lambda(\alpha)$ sets is j and the number of ${}^*\Lambda(\beta)$ sets is k, this proves (γ).

We now define the measure $\mu'(\Lambda')$ for sets of $\mathcal{B}'$ as follows: if $\Lambda' \in \mathcal{B}'$ and Λ is the strip based on Λ', then we define $\mu'(\Lambda')$ by $\mu'(\Lambda') = \mu(\Lambda)$.

LEMMA 7. *Both S_1 and S_1^{-1} take sets of $\mathcal{B}'$ into sets of $\mathcal{B}'$ and preserve μ'-measure.*

Proof. Because S_1 and S_1^{-1} take sets of $\mathcal{B}$ into sets of $\mathcal{B}$ and take Ω' into Ω', it follows that they take sets of $\mathcal{B}'$ into sets of $\mathcal{B}'$. Consequently, by Lemma 6, if Λ is a strip, then $\mu(S_1\Lambda) = \mu(\Lambda) = \mu(S_1^{-1}\Lambda)$ and hence μ'-measure is preserved by S_1 and S_1^{-1}.

LEMMA 8. *If Λ is a strip and a Borel set and $0 \leqq a < b \leqq 1$, and if $M = \Lambda \cdot \underset{P}{[}a < j(P) < b]$, then*

$$\mu(M) = (b - a)\mu(\Lambda).$$

Proof. Because of Lemma 6 it is sufficient to prove this assuming $b \leqq \frac{1}{2}$. Also it is sufficient to show that if N is any positive integer and if M_1 and M_2 are defined by

$$\begin{aligned} M_1 &= \Lambda \cdot \underset{P}{[}(j-1)2^{-N} < j(P) < j2^{-N}] \\ M_2 &= \Lambda \cdot \underset{P}{[}j2^{-N} < j(P) < (j+1)2^{-N}] \end{aligned} \qquad (j2^{-N} \leqq \tfrac{1}{2}),$$

then $\mu(M_1) = \mu(M_2)$.

We note that $g_{M_i}(P) = 2^{-N}g_R(P)$ for $P \in M_i$, so it will be sufficient to show that

$$\int_{M_1} \frac{1}{g_{M_1}(P)}\,dm = \int_{M_2} \frac{1}{g_{M_2}(P)}\,dm.$$

To show this we first define the sets ${}^nM_1^k$, ${}^nM_2^k$ by

$${}^nM_1^k = \underset{P}{[}k2^{-n} \leqq g_{M_1}(P) < (k+1)2^{-n}] \cdot \underset{P}{[}g_{M_1}(P) - f_{M_1}(P) < k2^{-n}],$$

$${}^nM_2^k = \underset{P}{[}k2^{-n} \leqq g_{M_2}(P) < (k+1)2^{-n}] \cdot \underset{P}{[}f_{M_2}(P) < k2^{-n}]$$

and define the functions $g_1^n(P)$, $g_2^n(P)$ by

$$g_1^n(P) = \begin{cases} k2^{-n} & \text{if } P \in {}^nM_1^k, \\ 0 & \text{if } P \in' \sum_k {}^nM_1^k, \end{cases}$$

$$g_2^n(P) = \begin{cases} k2^{-n} & \text{if } P \in {}^nM_2^k, \\ 0 & \text{if } P \in' \sum_k {}^nM_2^k. \end{cases}$$

Then we note (as in the proof of Lemma 6) that up to a set of measure 0 the sets ${}^nM_1^k$, ${}^nM_2^k$ are all translations of one another under members of T_t and hence

have the same m-measure; also $g^n_{M_i}(P) \to g_{M_i}(P)$ as $n \to \infty$ for $i = 1, 2$, and the $g^n_i(P)$ are uniformly bounded for large n. These facts show that

$$\int_{M_1} \frac{1}{g^n_1(P)}\, dm = \sum_k \frac{2^n}{k}\, m({}^nM^k_1) = \sum_k \frac{2^n}{k}\, m({}^nM^k_2) = \int_{M_2} \frac{1}{g^n_2(P)}\, dm$$

and

$$\int_{M_i} \frac{1}{g^n_i(P)}\, dm \to \int_{M_i} \frac{1}{g_{M_i}(P)}\, dm \qquad (i = 1, 2;\, n \to \infty)$$

and hence that

$$\int_{M_1} \frac{1}{g_{M_1}(P)}\, dm = \int_{M_2} \frac{1}{g_{M_2}(P)}\, dm.$$

THEOREM 4. *S_t is isomorphic to that flow on $\Omega' \times I$ which is built on the measure-preserving transformation S_1 on Ω'. The measure on Ω_1 that is carried over from the product measure on $\Omega' \times I$ is invariant under S_t and is equivalent to the original measure on Ω_1. S_t is a measurable flow.*

Proof. Denote by S'_t the flow on $\Omega' \times I$ which is built on S_1 on Ω'. We have remarked before that the 1:1 correspondence that we have set up between $\Omega' \times I$ and Ω_1 carries S'_t into S_t; i.e., if P corresponds to (P', x), then S_tP corresponds to $S'_t(P', x)$. Because the product measure on $\Omega' \times I$ is invariant under S'_t, it is also invariant under S_t; denote this measure on Ω_1 by $\bar{\mu}(M)$. We want now to show that $\bar{\mu}(M)$ is equivalent to $m(M)$, and it is for this purpose that we had to prove Lemmas 5, 6, 7 and 8.

Consider sets of the forms:

$$\Lambda \cdot \underset{P}{[}a < j(P) < b], \qquad \Lambda \cdot \underset{P}{[}a \leqq j(P) < b],$$

$$\Lambda \cdot \underset{P}{[}a \leqq j(P) \leqq b], \qquad \Lambda \cdot \underset{P}{[}a < j(P) \leqq b],$$

where Λ is a strip and a Borel set. Those sets which are finite disjunct sums of such sets form a field, $\mathcal{F}$, and the Borel field determined by this field is the collection of Borel sets, by Lemma 5. By Lemma 8 we see that if $M \in \mathcal{F}$ then

$$\bar{\mu}(M) = \int_M \frac{1}{g_{R^-}(P) + g_{CR^-}(P)}\, dm.$$

It follows from the fact that the normal family of sets determined by $\mathcal{F}$ is the same as the Borel field determined by $\mathcal{F}$ that if M is any Borel set then

$$\bar{\mu}(M) = \int_M \frac{1}{g_{R^-}(P) + g_{CR^-}(P)}\, dm.$$

Hence $\bar{\mu}$-measure is equivalent to m-measure.

Because S_1 and S_1^{-1} take Borel sets into Borel sets, it is easily seen that if Λ is a strip and a Borel set then the (t, P)-set

$$\underset{(t,P)}{[}\, S'_tP \in \Lambda \cdot \underset{P}{[}a < j(P) < b]]$$

is measurable in (t, P)-space; hence, since the sets of the form

$$\Lambda \cdot \underset{P}{[}a < j(P) < b]$$

determine the Borel sets in Ω_1, we see that if M is any Borel set in Ω_1 then

$$\underset{(t,P)}{[} S_t'P \in M]$$

is measurable. This proves that S_t' is a measurable flow. Because the measure invariant under S_t is equivalent to m-measure, and because we already know that S_t takes Borel sets into Borel sets, it follows that S_t takes measurable sets into measurable sets.

We have now proved our main theorem; we restate it as follows:

THEOREM 5. *If T_t is a continuous ergodic flow on a separable metric space of finite measure, then $\Omega = \Omega_1 + \Omega_2$, where*

(1) *Ω_1 and Ω_2 are disjunct invariant Borel sets and $m(\Omega_2) = 0$,*

(2) *by a change of velocities along the paths of T_t on Ω_1 it is possible to obtain a flow S_t which is isomorphic to a flow built on a measure-preserving transformation, and such that the measure invariant under S_t is equivalent to the original measure on Ω_1.*

As we pointed out before this theorem shows that by a change of velocities we can obtain a flow with an infinite number of eigenfunctions. This leads to the question whether it is possible by a change of velocities to obtain a flow with a pure point spectrum; i.e., a flow S_t for which there exist a sequence of numbers $\{\lambda_n\}$ and a complete orthogonal set of functions $\{\varphi_n(P)\}$ with the property that

$$\varphi_n(S_tP) = e^{i\lambda_n t}\varphi_n(P).$$

We shall not settle this question although it seems likely that the answer is negative; however, we shall prove the following theorem which is related to this matter.

THEOREM 6. *If T_t is a flow built on T, then T_t has a pure point spectrum if and only if T has a pure point spectrum.*[8]

Proof. We denote by Ω the space on which T is defined, and denote points of Ω by P. We denote points of $\Omega \times I$ on which T_t is defined by (P, x). As has been shown by von Neumann, $F(P, x)$ is an eigenfunction of T_t of eigenvalue λ if and only if[9]

$$F(P, x) = e^{i\lambda x}\psi(P),$$

where $\psi(P) \in L_2$ on Ω and satisfies $\psi(TP) = e^{i\lambda}\psi(P)$. Therefore the eigenfunctions of T_t consist of functions of the form

$$e^{i(\lambda_n+2\pi m)}\psi_n(P), \tag{i}$$

[8] T is said to have a pure point spectrum if there exist a sequence of numbers $\{\lambda_n\}$ and a complete orthogonal set of functions $\{\psi_n(P)\}$ with the property that $\psi_n(TP) = e^{i\lambda_n}\psi_n(P)$.

[9] See [6], p. 637. This property is stated in [6] only for a special transformation T, but the proof given there proves our statement.

where the function $\psi_n(P)$ is an eigenfunction of T of eigenvalue λ_n. It is easy to see that the functions (i) are complete in $\Omega \times I$ if and only if the functions $\psi_n(P)$ are complete in Ω.

We shall now prove a simple result about the form of a continuous ergodic flow; first, it is necessary to give the following definition.

DEFINITION 10. Let T be a measure-preserving transformation of a space Ω' into itself. Consider the product space $\Omega' \times T$ of Ω' with the real line, defining measure multiplicatively on this product space. Now let $F(P)$ be a positive function defined on Ω' and consider the subspace Ω of $\Omega' \times T$ consisting of points (P', x) for which $0 \leqq x < F(P')$. Define the flow T_t on Ω by

$$T_t(P', x) = (T^n P', x + t - F(P) - \cdots - F(T^n P))$$

if

$$0 \leqq F(P') + \cdots + F(T^n P') - x \leqq t < F(P') + \cdots + F(T^{n+1} P') - x,$$

$$T_t(P', x) = T_{-t}^{-1}(P', x) \quad \text{if} \quad t < 0.$$

We shall call this flow a *flow built under the function* $F(P)$ *on the measure-preserving transformation* T.

THEOREM 7. *If* T_t *is a continuous ergodic flow on a separable metric space of* Ω *finite measure, then* $\Omega = \Omega_1 + \Omega_2$, *where*

(1) Ω_1 *and* Ω_2 *are disjunct invariant Borel sets and* $m(\Omega_2) = 0$,

(2) T_t *on* Ω_1 *is isomorphic to a flow built under a function on a measure-preserving transformation.*

Proof. To prove this we take a regular set R and let Ω_1 and Ω_2 be the sets associated with R in the definition of a regular set. Then we again define Ω' as the set of points where the paths of T_t enter R, i.e., we define Ω' to be $\Omega_1(R^- - R)$. Then for $P' \epsilon \Omega'$ we define the function $k(P')$ by

$$k(P') = \text{minimum positive } t \text{ such that } P'_t \epsilon R^- - R.$$

Then with measure on Ω' taken as the μ'-measure defined above and T as the transformation used above on Ω' it is easy to see that T_t on Ω_1 is isomorphic to the flow built under the function $k(P')$ on the transformation T on Ω'.

BIBLIOGRAPHY

1. G. D. BIRKHOFF, *Proof of the ergodic theorem*, Proc. Nat. Acad. Sc., vol. 18(1932), pp. 656–660.
2. M. BEBOUTOFF AND W. STEPANOFF, *Sur la mesure invariante dans les systèmes dynamiques qui ne diffèrent que par le temps*, Recueil Math., vol. 7(1940), pp. 143–166.
3. E. HOPF, *Ergodentheorie*, Berlin, 1937.
4. B. O. KOOPMAN, *Hamiltonian systems and linear transformations in Hilbert space*, Proc. Nat. Acad. Sc., vol. 17(1931), pp. 315–318.
5. B. O. KOOPMAN AND J. v. NEUMANN, *Dynamical systems of continuous spectra*, Proc. Nat. Acad. Sc., vol. 18(1932), pp. 255–266.
6. J. v. NEUMANN, *Zur Operatorenmethode in der klassischen Mechanik*, Ann. of Math., vol. 33(1932), pp. 587–642.
7. N. WIENER, *The ergodic theorem*, this Journal, vol. 5(1939), pp. 1–18.

INSTITUTE FOR ADVANCED STUDY.

SUMS OF n-TH POWERS OF QUADRATIC INTEGERS

By Ivan Niven

1. **Introduction.** Consider the quadratic fields $R(\theta)$ defined by $\theta^2 = m$, m being a square-free rational integer not 0 or 1. Our first problem is to determine for which fields $R(\theta)$ every integer of the field is expressible as a sum of n-th powers of integers of the field. For odd powers this question is answered in §3 (Theorem 3), necessary and sufficient conditions being given in terms of m and n. For even powers, necessary and sufficient conditions are again given, but these are not as explicit as in the case of odd powers; this situation is treated in §6.

The second problem is to determine necessary and sufficient conditions that an integer of $R(\theta)$ be expressible as a sum of n-th powers of integers of the field. If n is odd, again we are able to give a complete answer; this is done in §4 (Theorems 4 and 5). For even powers we treat only imaginary fields, that is, fields for which m is negative; this is the material of §5 (Theorems 6, 7, and 8).

Quadratic integers are usually of the form $x + y\theta$, x and y being rational integers (as are all Roman letters in this paper). We shall say that such quadratic integers are of *the first kind*. When $m \equiv 1 \pmod 4$, however, quadratic integers may also be of the form $\frac{1}{2}(u + v\theta)$, u and v being odd. These will be called integers of *the second kind*. In §2 we study powers of quadratic integers of the second kind, and determine whether these powers are of the first or second kind. We show (in Theorems 1 and 2) that the n-th power of an integer of the second kind is an integer of the first kind if and only if n is divisible by 3 and m is congruent to 5 modulo 8. This result is used extensively throughout the paper, wherever integers of the second kind are under consideration.

2. **Powers of quadratic integers of the second kind.** We first prove a series of lemmas.

Lemma 1. *The product of two integers of the second kind,* $\frac{1}{2}(a + b\theta)$ *and* $\frac{1}{2}(c + d\theta)$, *is an integer of the second kind if and only if* $abcd \equiv 1 \pmod 4$.

Using the fact that m is congruent to 1 modulo 4, we see that the proof is an easy consequence of elementary congruence theory. Similarly we obtain the following result.

Lemma 2. *The product of an integer* $c + d\theta$ *of the first kind and an integer of the second kind is an integer of the second kind if and only if* c *and* d *are incongruent modulo* 2.

Lemma 3. *If an integer* $x + y\theta$ *has the property that* x *and* y *are incongruent modulo* 2, *then any power of this integer has the same property.*

Received November 20, 1940.

We write

$$(x + y\theta)^n = X + Y\theta,$$

so that

$$X = x^n + \binom{n}{2}x^{n-2}y^2m + \binom{n}{4}x^{n-4}y^4m^2 + \cdots ,$$

and

$$Y = \binom{n}{1}x^{n-1}y + \binom{n}{3}x^{n-3}y^3m + \binom{n}{5}x^{n-5}y^5m^2 + \cdots . \tag{1}$$

Since one of x and y is odd and the other even, we see that $X + Y$ is congruent to x^n or y^n modulo 2 according as x or y is odd, and this implies the result.

LEMMA 4. *Let m be congruent to 5 modulo 8. Then the cube of any integer of the second kind is an integer of the first kind.*

This follows from examination of the equation

$$(\tfrac{1}{2}a + \tfrac{1}{2}b\theta)^3 = \tfrac{1}{8}(a^3 + 3ab^2m) + \tfrac{1}{8}\theta(3a^2b + b^3m), \tag{2}$$

and the fact that a^2 is congruent to b^2 modulo 8, a and b being odd.

THEOREM 1. *Let m be congruent to 5 modulo 8. Then the n-th power of any integer of the second kind is of the first or second kind according as 3 is or is not a divisor of n.*

By Lemma 4 every cube is of the first kind; consequently any power of a cube is of the first kind.

Next we consider n-th powers where n is not a multiple of 3. First we show that the cube (2) has the property described in Lemma 3. That is, we prove that

$$\tfrac{1}{8}(a^3 + 3ab^2m) + \tfrac{1}{8}(3a^2b + b^3m) \equiv 1 \pmod 2,$$

or the equivalent congruence

$$a^3 + 3ab^2m + 3a^2b + b^3m \equiv 8 \pmod{16}. \tag{3}$$

It is easily seen that

$$a + 3b \equiv 0 \pmod 2 \qquad \text{and} \qquad a^2 + 3b^2m \equiv a^2 - b^2 \equiv 0 \pmod 8.$$

Multiplication of these two congruences gives

$$a^3 + 3a^2b + 3ab^2m + 9b^3m \equiv 0 \pmod{16}. \tag{4}$$

Since b and m are odd, $8b^3m$ is congruent to 8 modulo 16, and hence (4) implies (3).

We have proved that the cube of an integer of the second kind has integral coördinates which are incongruent modulo 2; Lemma 3 shows that the $(3k)$-th power of an integer of the second kind has the same property. Hence by Lemma

2 the $(3k + 1)$-th power of an integer of the second kind is of the second kind. The same argument applies to $(3k + 2)$-th powers, because Lemma 1 implies that the square of an integer of the second kind is of the second kind. This completes the proof of the theorem.

THEOREM 2. *Let m be congruent to* 1 *modulo* 8. *Then any power of an integer of the second kind is an integer of the second kind.*

We prove this by mathematical induction. Assume that

$$(\tfrac{1}{2}a + \tfrac{1}{2}b\theta)^{n-1} = \tfrac{1}{2}c + \tfrac{1}{2}d\theta, \qquad abcd \equiv 1 \pmod 4.$$

Note that this is true if $n = 2$. Then we have

$$(\tfrac{1}{2}a + \tfrac{1}{2}b\theta)^{n} = \tfrac{1}{4}(ac + bdm) + \tfrac{1}{4}\theta(ad + bc).$$

By Lemma 1 this is an integer of the second kind. In order to complete the proof we must show that

$$ab\,\frac{ac + bdm}{2}\cdot\frac{ad + bc}{2} \equiv 1 \pmod 4. \tag{5}$$

First let ab be congruent to 1 modulo 4. Then congruence (5) is equivalent to

$$\tfrac{1}{2}(ac + bdm) \equiv \tfrac{1}{2}(ad + bc) \pmod 4,$$

or

$$ac + bd - ad - bc \equiv 0 \pmod 8, \tag{6}$$

since m is congruent to 1 modulo 8. The truth of this congruence follows from factoring the left side and observing that $a - b$ is divisible by 4 and $c - d$ is even.

In the second place, let ab be congruent to 3 modulo 4. Then congruence (5) is now equivalent to

$$ac + bd - ad - bc \equiv 4 \pmod 8. \tag{7}$$

Now the hypothesis that $abcd \equiv 1 \pmod 4$ implies that $cd \equiv 3 \pmod 4$. Hence we can conclude that both $a - b$ and $c - d$ are even but not divisible by 4, and this establishes (7).

3. Fields whose integers are sums of odd powers. In this section we determine those fields all of whose integers are sums of n-th powers, n being odd.

THEOREM 3. *Let n be any odd integer greater than* 1. *Then every integer of the quadratic field $R(\theta)$, defined by $\theta^2 = m$, is expressible as a sum of n-th powers of integers of the field if and only if $(m, n) = 1$ and, when $3 \mid n$, m is not congruent to* 5 *modulo* 8.

That the latter condition is necessary follows from Theorem 1. For if every n-th power is an integer of the first kind, so is every sum of n-th powers.

In order to prove that the condition $(m, n) = 1$ is necessary, we examine $(x + y\theta)^n$. The coefficient of θ in this n-th power, by equation (1), is a sum of terms each of which is divisible by the greatest common divisor of m and n. It is clear that the same result holds for n-th powers of integers of the second kind, because (m, n) is an odd integer. Hence if $(m, n) > 1$, the integer $a + b\theta$ is not expressible as a sum of n-th powers if b is not divisible by (m, n).

We now show that the conditions are sufficient. First we treat integers of the first kind. The method was suggested by a paper of Tornheim.[1] Define D_n by the equation

$$(8) \qquad D_n = \binom{n}{1}\binom{n}{2}\cdots\binom{n}{n-1}\begin{vmatrix} 1 & 1^2 & \cdots & 1^{n-1} \\ 2 & 2^2 & \cdots & 2^{n-1} \\ \cdots & \cdots & \cdots & \cdots \\ n-1 & (n-1)^2 & \cdots & (n-1)^{n-1} \end{vmatrix}.$$

By the Euclidean algorithm for rational integers, any quadratic integer $a + b\theta$ satisfies an equation of the form

$$(9) \qquad a + b\theta = \gamma D_n + \gamma_1 \qquad (\gamma_1 = c + d\theta, \quad 0 \leqq c < D_n, \quad 0 \leqq d < D_n),$$

where γ is a quadratic integer. Now consider the equations

$$(10) \qquad (\gamma + r)^n = \gamma^n + \binom{n}{1}\gamma^{n-1}r + \cdots + r^n \qquad (r = 1, 2, \cdots, n-1).$$

We can consider these to be $n - 1$ linear equations in the $n - 1$ unknowns $\gamma, \gamma^2, \cdots, \gamma^{n-1}$, and solve for γ by Cramer's rule. Thus we have

$$(11) \qquad D_n\gamma = E_n,$$

where E_n is a linear homogeneous polynomial in the n-th powers γ^n, r^n, and $(\gamma + r)^n$, with r ranging over the values $1, 2, \cdots, n - 1$. If any of the integral coefficients of this polynomial are negative, they can be made positive by incorporating the signs into the n-th powers, n being odd. Having thus shown that $D_n\gamma$ is a sum of n-th powers, we turn to the quantity γ_1 of equation (9). We write

$$(12) \qquad \theta^n = h_1\theta, \qquad (1 + \theta)^n = h_2\theta + l_2,$$

where

$$(13) \qquad h_1 = m^{\frac{1}{2}(n-1)}, \qquad h_2 = n + \binom{n}{3}m + \binom{n}{5}m^2 + \cdots.$$

The integers h_1 and h_2 are prime to each other; for if a prime p divides h_1 it must divide m and consequently $h_2 - n$, but p does not divide n and hence not h_2.

[1] Leonard Tornheim, *Sums of n-th powers in fields of prime characteristic*, this Journal, vol. 4(1938), p. 360.

Then there exist integers u and v such that $uh_1 + vh_2 = 1$, and it follows that

$$u\theta^n + v(1 + \theta)^n = \theta + vl_2,$$

and

$$\gamma_1 = c + d\theta = ud\theta^n + vd(1 + \theta)^n + c - vdl_2.$$

Again incorporating any negative signs into the n-th powers, and writing $c - vdl_2$ as a sum of n-th powers of 1 or -1, we have expressed γ_1 as a sum of n-th powers.

Finally we must consider any integer of the second kind, $\frac{1}{2}(a + b\theta)$. Theorem 1 shows that $(\frac{1}{2} + \frac{1}{2}\theta)^n$ is an integer of the second kind. Hence the integer

$$\frac{1}{2}(a + b\theta) - (\tfrac{1}{2} + \tfrac{1}{2}\theta)^n \tag{14}$$

is of the first kind, and consequently is expressible as a sum of n-th powers. So also is $\frac{1}{2}(a + b\theta)$, and this completes the proof of the theorem.

4. **Integers which are sums of odd powers.** We now determine necessary and sufficient conditions that any integer be expressible as a sum of n-th powers. We begin by treating integers of the first kind.

THEOREM 4. *Let n be any odd integer greater than* 1, *and define k by the equation*

$$k = (m^{\frac{1}{2}(n-1)}, n). \tag{15}$$

Then $a + b\theta$ is a sum of n-th powers of integers of the field $R(\theta)$ if and only if $3k \mid b$ when $3 \mid k$ and $m \equiv 3 \pmod 9$ and $n > 3$, and $k \mid b$ otherwise.

We prove that this condition is necessary by examining equation (1). It is clear that $k \mid Y$, because k divides each of the coefficients

$$\binom{n}{1}, \quad \binom{n}{3}m, \quad \binom{n}{5}m^2, \quad \cdots . \tag{16}$$

Now let 3^α, where $\alpha > 0$, be the highest power of 3 dividing k, and suppose $m \equiv 3 \pmod 9$. Then $3^{\alpha+1}$ divides all but the first two terms of the series (16). Thus we can show that $3^{\alpha+1}$ divides Y of equation (1) by showing that $3^{\alpha+1}$ divides the first two terms of the expression for Y, and these terms may be written in the form

$$\tfrac{1}{2}nx^{n-3}y\left\{2x^2 - (n-1)(n-2)y^2\cdot\frac{m}{3}\right\}. \tag{17}$$

If x or y is divisible by 3, then this expression is divisible by $3^{\alpha+1}$, since 3^α divides n. If, on the other hand, x and y are prime to 3, then x^2 and y^2 are congruent to 1 modulo 3. Hence we have

$$2x^2 - (n-1)(n-2)y^2\cdot\frac{m}{3} \equiv 2 - (-1)(-2)(1)(1) \equiv 0 \pmod 3,$$

so that expression (17) is divisible by $3^{\alpha+1}$. An exception must be made to this argument when we take $x = 0$, y prime to 3, and $n = 3$. Hence the condition $n > 3$ in the theorem.

We have shown that Y of equation (1) is divisible by $3k$ or k in the different cases. Since k is an odd integer, these arguments hold for n-th powers of integers of the second kind, and consequently any sum of n-th powers has this divisibility property.

Conversely, suppose that b is divisible by $3k$ or k as outlined in the theorem. Equation (9) can again be written, and $D_n\gamma$ is a sum of n-th powers. Since D_n is divisible by $3k$, it follows that d is divisible by $3k$ or k according as b is so divisible.

In order to prove that $c + d\theta$ is a sum of n-th powers, we use equations (12), (13), and

$$(18) \qquad (1 + 2\theta)^n = h_3\theta + l_3, \qquad h_3 = 2n + 8\binom{n}{3}m + 32\binom{n}{5}m^2 + \cdots.$$

Case 1. $3 \nmid k$. We show that $(h_1, h_2) = k$. First, if a prime does not divide k, then it does not divide h_1. Second, if p^α, with $\alpha > 0$, is the highest power of a prime p which divides k, it is obvious from (13) that p^α divides h_1 and h_2. Furthermore, since $p > 3$, $p^{\alpha+1}$ divides $h_2 - n$ but not n, so it does not divide h_2.

Hence there exist rational integers u and v such that $uh_1 + vh_2 = k$. And if $d = kd_1$ we have

$$c + d\theta = d_1u\theta^n + d_1v(1 + \theta)^n + c - d_1vl_2,$$

the right side of this equation being a sum of n-th powers.

Case 2. $3 \mid k$, $m \equiv 6 \pmod 9$. Again we assume that $k \mid d$, and we show that $(h_1, h_2) = k$. For primes greater than 3 the proof of this is the same as in Case 1. Next let 3^α be the highest power of 3 dividing k, whence 3^α divides n, but $3^{\alpha+1}$ does not. The second of equations (13) shows that $3^{\alpha+1}$ divides

$$h_2 - n - \binom{n}{3}m.$$

We write

$$n + \binom{n}{3}m = \tfrac{1}{2}n\left\{2 - (n-1)(n-2)\frac{m}{3}\right\},$$

and

$$2 - (n-1)(n-2)\frac{m}{3} \equiv 2 - (-1)(-2)(2) \equiv 1 \pmod 3.$$

This proves that 3^α is the highest power of 3 dividing $n + \binom{n}{3}m$, and hence the same result holds for h_2. Thus the g.c.d. of h_1 and h_2 is k, and the remainder of the proof is analogous to Case 1.

Case 3. $3 \mid k$, $m \equiv 3 \pmod 9$. If $3 \mid k$, then $3 \mid m$, and m must be congruent to 3 or 6 modulo 9, since it is square-free; thus Cases 1, 2, and 3 exhaust the possibilities. When $n = 3$, we have $\theta^n = m\theta$; hence $(h_1, h_2) = k = 3$ and the proof in Case 1 applies again. When $n > 3$, we assume that $3k \mid d$, this being in accord with the statement of the theorem. We show that $(h_1, h_2, h_3) = 3k$. For primes greater than 3, the proof is analogous to Case 1. Again let 3^α be the highest power of 3 which divides k. Now $3^{\alpha+2}$ divides all terms of the series (16) except the first two, so that by (13) and (18), $3^{\alpha+2}$ divides both

$$h_2 - n - \binom{n}{3}m \quad \text{and} \quad h_3 - 2n - 8\binom{n}{3}m.$$

But $3^{\alpha+1}$ is the highest power of 3 dividing both

$$n + \binom{n}{3}m \quad \text{and} \quad n + 4\binom{n}{3}m.$$

For the difference between these terms is

$$3\binom{n}{3}m = \tfrac{1}{2}n(n-1)(n-2)m,$$

which is divisible by $3^{\alpha+1}$ but not $3^{\alpha+2}$; also the congruence in Case 2 now reads

$$2 - (n-1)(n-2)\frac{m}{3} \equiv 2 - (-1)(-2)(1) \equiv 0 \pmod 3,$$

so that $3^{\alpha+1}$ divides $n + \binom{n}{3}m$ in the present case. Hence the highest power of 3 dividing h_2 and h_3 is $3^{\alpha+1}$. Noting that $3^{\alpha+1}$ divides h_1, we have proved that the g.c.d. of h_1, h_2, and h_3 is $3k$. And this enables us to find integers u, v, and w such that

$$uh_1 + vh_2 + wh_3 = 3k,$$

so that $c + d\theta$ is expressible as a linear combination of 1^n, θ^n, $(1 + \theta)^n$, and $(1 + 2\theta)^n$.

For integers of the second kind we can make an argument similar to that at the end of §3. The coefficient of θ in expression (14) is divisible by $3k$ or k according as b is so divisible. Hence Theorem 4 enables us to write the following result.

THEOREM 5. *Let n be any odd integer greater than* 1. *Define k by equation* (15). *Then an integer of the second kind, $\frac{1}{2}(a + b\theta)$, is expressible as a sum of n-th powers of integers of $R(\theta)$ if and only if b satisfies the conditions of Theorem* 4 *and, when $3 \mid n$, m is not congruent to* 5 *modulo* 8.

5. Integers which are sums of even powers. We restrict our discussion to imaginary fields, that is, fields for which m is negative. The first of equations (12) was used considerably in the last two sections, but is no longer true, m being

even. We consider $(x + y\theta)^n$, and particularly equation (1). Hensel[2] has shown that the g.c.d. of the infinite set of values Y (obtained by giving x and y all possible integral values) equals the g.c.d. of the finite set of values of Y obtained by assigning n consecutive integral values to x and n consecutive values to y. Let k denote the g.c.d. of the values of Y when, say, x and y range over the values $1, 2, \cdots, n$, that is,

$$k = (Y_{11}, Y_{12}, \cdots, Y_{1n}, Y_{21}, \cdots, Y_{nn}), \tag{19}$$

where Y_{ij} is the value of Y when $x = i$ and $y = j$. By Hensel's result, if $a + b\theta$ is expressible as a sum of n-th powers of integers of the first kind, then b is divisible by k.

THEOREM 6. *Let n be a positive even integer, and let m be a negative integer incongruent to* 1 *modulo* 4. *Then $a + b\theta$ is expressible as a sum of n-th powers of integers of $R(\theta)$ if and only if b is divisible by k as defined in* (19).

Having shown this condition to be necessary, we now assume that $k \mid b$, and show that $a + b\theta$ is a sum of n-th powers. Let $n = 2^t N$, where N is odd. Hilbert[3] has shown that to any positive integer s there corresponds the identity

$$(x_1^2 + \cdots + x_5^2)^s = \sum_{h=1}^{M} r_h(a_{1h}x_1 + \cdots + a_{5h}x_5)^{2s}, \tag{20}$$

where the r_h are positive rational numbers and the a_{ih} are rational integers, and

$$M = \frac{(2s + 1) \cdots (2s + 4)}{1 \cdot 2 \cdot 3 \cdot 4}.$$

Thus if $u(s)$ denotes the g.c.d. of the denominators of the r_h, then we obtain from (20)

$$u(s)(x_1^2 + \cdots + x_5^2)^s = \sum_{h=1}^{M} R_h(a_{1h}x_1 + \cdots + a_{5h}x_5)^{2s}, \tag{21}$$

where the R_h are positive rational integers. Now define D by the equation

$$D = 2^r D_N\, u(N)u(2N)u(4N) \cdots u(2^{t-1}N), \tag{22}$$

where $r = N + 2N + 4N + \cdots + 2^{t-1}N$.

We now prove that $D\gamma$, where γ is any integer of $R(\theta)$, is a sum of n-th powers. First, by the proof of Theorem 3, $D_N\gamma$ is a sum of N-th powers:

$$D_N\gamma = \sum (x_i + y_i\theta)^N,$$

[2] K. Hensel, *Ueber den grössten gemeinsamen Theiler aller Zahlen, welche durch eine ganze Function von n Veränderlichen darstellbar sind*, Journal für die reine und angewandte Mathematik, vol. 116(1896), pp. 350–356.

[3] David Hilbert, *Lösung der Waringschen Problems*, Mathematische Annalen, vol. 67(1909), p. 283.

so that

$$2^N D_N \gamma = \sum (2x_i + 2y_i\theta)^N.$$

It has been shown by the author[4] that $2x_i + 2y_i\theta$ is expressible as a sum of three squares of integers of $R(\theta)$. Using this and Hilbert's identity (21) we obtain

$$2^N D_N \gamma u(N) = \sum (w_i + z_i\theta)^{2N}.$$

In the same way we can show that $2^N D_N \gamma u(N) 2^{2N} u(2N)$ is a sum of $(4N)$-th powers of integers of $R(\theta)$. Proceeding thus, we obtain the desired result.

Let $b = DB + d$, where $0 \leqq d < D$. By the necessity of the condition of the present theorem, the fact that $D\gamma$ is expressible as a sum of n-th powers implies that $k \mid D$; and since b is divisible by k by hypothesis, so also is d. Since k is the g.c.d. of the n^2 numbers Y_{ij}, it is expressible as a linear combination of these numbers with integral coefficients:

$$k = a_{11}Y_{11} + a_{12}Y_{12} + \cdots + a_{nn}Y_{nn}.$$

Recalling the definition of Y_{ij} in equation (19), we have

$$k\theta = a_{11}(1 + \theta)^n + a_{12}(1 + 2\theta)^n + \cdots + a_{nn}(n + n\theta)^n + q, \tag{23}$$

where q is a rational integer, positive or negative. Some of the coefficients a_{ij} in (23) may be negative. For example, suppose that a_{11} is negative. Then the coefficient of θ in $| a_{11} | \cdot (1 - \theta)^n$ is the same as that in $a_{11}(1 + \theta)^n$. Thus equation (23) can be replaced by

$$k\theta = | a_{11} | \cdot (1 \pm \theta)^n + | a_{12} | \cdot (1 \pm 2\theta)^n + \cdots + | a_{nn} | \cdot (n \pm n\theta)^n + q_1, \tag{24}$$

where the sign in each $(i \pm j\theta)$ is the same as the sign of a_{ij}. Note that the integer q has been replaced by q_1. Let $d = kd_1$. Then equation (24) enables us to write

$$\begin{aligned} a + b\theta &= a + DB\theta + d\theta \\ &= a + DB\theta + d_1 \cdot \sum_{i,j=1}^{n} | a_{ij} | \cdot (i \pm j\theta)^n + d_1 q_1. \end{aligned}$$

By the Euclidean algorithm there exist integers A and c such that $a + d_1 q_1 = AD + c$, where $0 \leqq c < D$. Hence we have

$$a + b\theta = D(A + B\theta) + c + d_1 \cdot \sum_{i,j=1}^{n} | a_{ij} | \cdot (i \pm j\theta)^n.$$

We have proved that $D(A + B\theta)$ is a sum of n-th powers; and the positive integer c is a sum of n-th powers of 1. Our theorem has now been established.

This theorem also applies in case $m \equiv 1 \pmod 4$, but it is concerned with integers of the first kind only. Theorem 1 gives

THEOREM 7. *Let n be a multiple of* 6, *and let m be a negative integer congruent to* 5 *modulo* 8. *Then $a + b\theta$ is expressible as a sum of n-th powers of integers of $R(\theta)$ if and only if b is divisible by k as defined in equation* (19).

[4] Ivan Niven, *Integers of quadratic fields as sums of squares*, Transactions of the American Mathematical Society, vol. 48(1940), p. 413.

The necessity of this condition depends upon the fact that the coördinate of θ in the n-th power of any integer of the second kind is divisible by as high a power of 2 as the coördinate of θ in, e.g., $(1 + 2\theta)^n$. The demonstration of this is not trivial, but a congruence proof similar to those in §2 is all that is needed.

Finally we turn to those imaginary fields wherein the n-th power of an integer of the second kind is of the second kind. We can write

$$(\tfrac{1}{2} + \tfrac{1}{2}\theta)^n = \tfrac{1}{2}(e + f\theta),$$

where e and f are odd. Let 2^β be the highest power of 2 dividing k of equation (19), and let $k = 2^\beta K$. It is clear that $K \mid f$, and moreover, if $\frac{1}{2}(a + b\theta)$ is a sum of n-th powers, then $K \mid b$.

Conversely, suppose that $K \mid b$. Let x be the least positive integer such that

$$b - xf \equiv 0 \pmod{2^{\beta+1}}.$$

Such an integer exists, because if x ranges over a complete residue system modulo $2^{\beta+1}$, so does xf. Therefore we have

$$\tfrac{1}{2}(a + b\theta) - x(\tfrac{1}{2} + \tfrac{1}{2}\theta)^n = \tfrac{1}{2}(a - xe) + \tfrac{1}{2}\theta(b - xf). \tag{25}$$

Now $\frac{1}{2}(b - xf)$ is divisible by 2^β and by K, and hence by k. The integer x is odd, so that $\frac{1}{2}(a - xe)$ is a rational integer. By the proof of Theorem 6, the integer on the right side of equation (25) is a sum of n-th powers. Hence $\frac{1}{2}(a + b\theta)$ is a sum of n-th powers.

THEOREM 8. *Let m be a negative integer congruent to* 1 *modulo* 4, *and let n be a positive even integer such that* $3 \nmid n$ *when* $m \equiv 5 \pmod 8$. *Define K as the odd integer obtained from k of equation* (19) *by dividing out the highest power of* 2. *Then an integer of $R(\theta)$ is expressible as a sum of n-th powers of integers of the field if and only if its imaginary coördinate is divisible by K.*

We have yet to prove this statement for integers of the first kind. The necessity of the condition is apparent. To prove it sufficient, suppose $K \mid b$, and write

$$a + b\theta - (\tfrac{1}{2} + \tfrac{1}{2}\theta)^n = \tfrac{1}{2}(2a - e) + \tfrac{1}{2}\theta(2b - f).$$

Since $K \mid 2b - f$, the integer of the second kind on the right side of this equation is a sum of n-th powers; so also, therefore, is $a + b\theta$.

6. **Fields whose integers are sums of even powers.** We are now in a position to decide whether or not all the integers of $R(\theta)$ are expressible as sums of even powers. Unfortunately the conditions are not given in terms of m and n explicitly, but must also involve the integer k of equation (19), which is a function of m and n.

THEOREM 9. *Let n be a positive even integer. Then every integer of $R(\theta)$ defined by $\theta^2 = m$ is expressible as a sum of n-th powers of integers of the field if*

and only if m is negative and congruent to 1 *modulo* 4, $m \equiv 1 \pmod 8$ *when* $3 \mid n$, *and the integer k of equation* (19) *is a power of* 2.

Theorem 8 shows that the conditions are sufficient. The last two conditions are seen to be necessary by Theorem 1 and the definition of k, respectively. It is necessary that m be negative. For when m is positive the integers of the field are real numbers, and consequently no negative integer is expressible as a sum of even powers.

Finally it is necessary that m be congruent to 1 modulo 4. For otherwise the integers of the field are of the first kind only. And since the square of any such integer has an even imaginary coördinate, any sum of even powers has the same property.

UNIVERSITY OF ILLINOIS.

EXTENSION OF HOMEOMORPHISMS INTO EUCLIDEAN AND HILBERT PARALLELOTOPES

By Ralph H. Fox

Let A be a closed subset of a space X, and let f be a homeomorphism of A into a space Y. A homeomorphism f^* of X into Y is called an *extension* of f when $f^*(x) = f(x)$ for every $x \epsilon A$. If A is the null set, every homeomorphism f^* of X into Y is an extension of f. Thus the problem of extending a given homeomorphism is a generalization of the problem of imbedding. Guided by this remark I shall prove the following generalization of the Menger-Nöbeling[1] imbedding theorem:

(1) *Let A be a compact subset of a separable metrizable space X. Let n be the dimension of $X - A$, and let y be a point of the n-dimensional parallelotope[2] E^n. If f is a homeomorphism of A into the $(q + n)$-dimensional parallelotope $Y = E^q \times E^n$, where $q \geqq 1 + \dim X$, and if $f(A) \subset E^q \times [y]$, then f can be extended to a homeomorphism of X into Y.*

The theorem holds also in the case $n = \infty$; it is then a generalization of the Urysohn imbedding theorem[3] and reads as follows:

(2) *Let A be a compact subset of a separable metrizable space X and let y be a point of the Hilbert parallelotope[2] E^∞. Any homeomorphism f of A into the Hilbert parallelotope $Y = E^\infty \times E^\infty$, such that $f(A) \subset E^\infty \times [y]$, can be extended to a homeomorphism of X into Y.*

For finite q every compact subset of the Euclidean q-dimensional space R^q is contained in a homeomorph of E^q; hence E^q may be replaced by R^q in the statement of (1). In the Menger-Nöbeling theorem it is immaterial whether E^q or R^q is used, but for (1) the use of E^q results in a theorem which is *a priori* stronger.

It is interesting to compare (1) with the theorems[4] of Gehman and Adkisson-

Received January 17, 1941.

[1] W. Hurewicz, *Über Abbildungen von endlichdimensionalen Räumen auf Teilmengen Cartesischer Räume*, Sitzungsberichte Preuss. Akad. Wiss., vol. 24(1933), pp. 754–768, where further references may be found. Also C. Kuratowski, *Sur les théorèmes du "plongement" dans la théorie de la dimension*, Fundamenta Mathematicae, vol. 28(1937), pp. 336–342. Some of the methods of the present note stem from this latter paper.

[2] The n-dimensional parallelotope E^n is the product of the closed interval [0, 1] with itself n times. The Hilbert parallelotope E^∞ is the product of [0, 1] with itself a countable number of times. We consider E^n as a subset of Euclidean n-space R^n in the usual way, and E^∞ as a subset of Hilbert space R^∞.

[3] Alexandroff and Hopf, *Topologie*, Berlin, 1935, p. 81.

[4] H. M. Gehman, *On extending a continuous (1-1) correspondence of two plane continuous curves to a correspondence of their planes*, Transactions of the American Mathematical Society, vol. 28(1926), pp. 252–265; V. W. Adkisson and Saunders MacLane, *Extending maps of plane Peano continua*, this Journal, vol. 6(1940), pp. 216–228, Theorem 2, where further references may be found. Cf. also G. Choquet, *Etude des homéomorphies planes*, Paris Comptes Rendus, vol. 206(1938), pp. 159–161.

MacLane. Although their theorems give a complete solution to the homeomorphism-extension problem: (E) *Under what conditions can a homeomorphism of a compact subset A of a space X into a Euclidean space R be extended to a homeomorphism of X into R?* when $X = R$ is the Euclidean plane and A is connected and locally connected, their methods lean heavily on the properties of the plane. In fact when $X = R$ is the Euclidean 3-space, the problem becomes intricately tangled with the well-known knot problem, so that a complete solution could scarcely be expected in this case.

In contrast, the conditions of theorem (1) impose no bound on the dimension of R, but furnish a solution of (E) only when suitable restrictions have been placed on the homeomorphism. Thus (1) implies that every homeomorphism of a compact subset of a 1-dimensional space into the plane can be extended to a homeomorphism into 3-space. Of course the more general statements below guarantee much more than the mere existence of a homeomorphism. That the homeomorphic extensions of a given homeomorphism are proved to constitute a residual set in the space of all extensions explains why the conditions of theorem (1) are so restrictive from the standpoint of existence only.

Denote by Y^X the space whose elements are the continuous mappings of a separable metric space X into a bounded metric space Y, metrized by the formula $d(g_1, g_2) = \sup_{x \epsilon X} d(g_1(x), g_2(x))$. Let f be a homeomorphism of a compact subset A of X into Y. For any $\epsilon > 0$ denote by $N_\epsilon = N_\epsilon(A)$ the open ϵ-neighborhood of A in X and let $K_\epsilon = A + (X - N_\epsilon)$.

Let $\mathfrak{F}$ denote the set of continuous mappings of X into Y which are extensions of f. For any pair of disjoint closed sets C_1 and C_2 of X let $\mathfrak{F}(C_1, C_2)$ denote the set of elements g of F such that $\overline{g(C_1)} \cdot \overline{g(C_2)} = 0$; for any $\epsilon > 0$ let $\mathfrak{F}_\epsilon(C_1, C_2) = \mathfrak{F}(C_1 \cdot K_\epsilon, C_2 \cdot K_\epsilon)$.

(3) *For any disjoint sets C_1 and C_2 of X we have* $\mathfrak{F}(C_1, C_2) = \prod_{k=1}^{\infty} \mathfrak{F}_{1/k}(C_1, C_2)$.

Obviously $\mathfrak{F}(C_1, C_2) \subset \prod_{k=1}^{\infty} \mathfrak{F}_{1/k}(C_1, C_2)$. Therefore we need only prove that $\mathfrak{F}(C_1, C_2) \supset \prod_{k=1}^{\infty} \mathfrak{F}_{1/k}(C_1, C_2)$. Suppose $g \epsilon \prod_{k=1}^{\infty} \mathfrak{F}_{1/k}(C_1, C_2) - \mathfrak{F}(C_1, C_2)$ and let $y \epsilon \overline{g(C_1)} \cdot \overline{g(C_2)}$. Then there exist sequences $\{x_m^1\} \subset C_1$ and $\{x_n^2\} \subset C_2$ such that $g(x_m^1) \to y$ and $g(x_n^2) \to y$. Since, by hypothesis, $y \epsilon' \overline{g(C_1 \cdot K_{1/k})} \cdot \overline{g(C_2 \cdot K_{1/k})}$ for any k, it follows that $\sum x_m^1 + \sum x_n^2$ intersects each of the open sets $X - K_{1/k} = N_{1/k} - A$. Therefore at least one of the sets $\sum x_m^1$ or $\sum x_n^2$, say $\sum x_m^1$, intersects each of the sets $N_{1/k} - A$. It follows[5] that a subsequence of $\{x_m^1\}$ converges to a point x^1 of A. For simplicity let us assume $x_m^1 \to x^1$. Then $g(x_m^1) \to g(x^1)$, hence $g(x^1) = y$. Since, therefore, $y \epsilon \overline{g(C_1 \cdot K_{1/k})}$ for every k, it follows that $y \epsilon' \overline{g(C_2 \cdot K_{1/k})}$ for any k. Hence $\sum x_n^2$ intersects each of the sets $N_{1/k} - A$ and hence a subsequence of $\{x_n^2\}$ converges to a point x^2

[5] Alexandroff and Hopf, loc. cit., p. 100, Satz II.

of A. As before, $g(x^2) = y$. But $g \mid A$ is a homeomorphism so that $x^1 = x^2$. This implies that $C_1 \cdot C_2 \neq 0$, and this is contrary to hypothesis.

The main theorem (1) is a corollary of the more general theorem

(4) *Let A be a compact subset of a separable metric space X, and let n be the dimension of $X - A$. Let f be a homeomorphism of A into the p-dimensional parallelotope $Y = E^p \subset R^p$, where $p \geqq 2n + 1$.*

If, for every linear subspace L of R^p of dimension less than n which is disjoint to $f(A)$, the complement of the span[6] *of $f(A)$ and L is dense in Y, then the set $\mathfrak{H}$ of topological extensions of f is a residual set in the complete metric space $\mathfrak{F}$ of continuous extensions of f.*

Proof of (4). Since $\mathfrak{F}$ is obviously closed in Y^X and since Y is compact, so that Y^X is complete, it follows that $\mathfrak{F}$ is complete. Hence,[7] in order to prove that $\mathfrak{H}$ is residual, it is sufficient to show that $\mathfrak{F}(C_1, C_2)$ is dense and open in $\mathfrak{F}$. Since[8] C_1 and C_2 are closed and disjoint, $\mathfrak{F}(C_1, C_2)$ is open in $\mathfrak{F}$. Since $\mathfrak{F}_\epsilon(C_1, C_2) \subset \mathfrak{F}_\delta(C_1, C_2)$ whenever $\epsilon \leqq \delta$, it follows from (3) that it is sufficient to show that for every $f^* \epsilon \mathfrak{F}$ and $\epsilon > 0$ an element $g \epsilon \mathfrak{F}_\epsilon(C_1, C_2)$ can be found such that $d(f^*, g) < \epsilon$.

Let $f^* \epsilon \mathfrak{F}$ and let $\{G_1, G_2, \cdots, G_r\}$ be a finite open covering of $X - N_\epsilon = K_\epsilon - A$ which has the following properties:[9]

(5) $G_i \subset X - A$,

(6) $d(f^*(G_i)) < \frac{1}{2}\epsilon$,

(7) no G_i intersects both C_1 and C_2,

(8) the nerve of $\{G_i\}$ is at most n-dimensional.

One can now find r points $y_1, y_2, \cdots, y_r$ in Y such that

(9) $d(y_i, f^*(G_i)) < \frac{1}{2}\epsilon$,

(10) for no $m < 2n + 3$ are m of the points of $\{y_i\}$ contained in a linear subspace of R^p of dimension $m - 2$, i.e., any set of fewer than $2n + 3$ of the points of $\{y_i\}$ are independent,

(11) the linear subspace of R^p determined by any $k < n + 2$ of the points of $\{y_i\}$ is disjoint to $f(A)$.

In fact the points $y_1, \cdots, y_r$ are picked out inductively. Choose for y_1 any point satisfying (9) and lying in $Y - f(A)$; this is possible since by hypothesis $Y - f(A)$ is dense in Y. Suppose $y_1, \cdots, y_t$ $(t < r)$ have been chosen to satisfy (9), (10), and (11). The linear space determined by any set of $m - 1$ of the points of $y_1, \cdots, y_t$ is of dimension less than $2n + 1$. Hence the union of all such linear spaces is of dimension less than $2n + 1$. By hypothesis the span of $f(A)$ and the linear space determined by any set of $k - 1$ of the points of $y_1, \cdots, y_t$ has a dense complement in Y. Thus there exist points in the

[6] By the *span* of two disjoint subsets B_1 and B_2 of a linear space R is meant here the union of the lines of R which intersect both B_1 and B_2.

[7] Kuratowski, loc. cit., p. 339.

[8] Ibid., p. 338, footnote 1.

[9] Ibid., p. 337.

$\frac{1}{2}\epsilon$-neighborhood of $f^*(G_i)$ in Y which do not lie on any of the linear spaces determined by a set of $m - 1$ points of $y_1, \cdots, y_t$ or in the span of $f(A)$ with any of the linear spaces determined by $k - 1$ points of $y_1, \cdots, y_t$. Choose y_{t+1} one of these points; it is easily verified that $y_1, \cdots, y_{t+1}$ satisfies (9), (10), and (11), and the induction is complete.

Let

(12) $\rho_i(x) = d(x, X - G_i)$ $(i = 1, \cdots, r)$,

and let $\theta = \theta_\epsilon$ be a mapping of X into $[0, 1]$ such that

(13) $\theta(X - \sum G_i) = 1$,

(14) $\theta(X - N_\epsilon) = 0$.

Let $g(x)$ be the center of gravity of the system which is obtained by placing a weight $\rho_i(x)$ at the point y_i $(i = 1, \cdots, r)$ and a weight $\theta(x)$ at the point $f^*(x)$.

Explicitly

$$g(x) = \frac{\theta(x)f^*(x) + \sum \rho_i(x)y_i}{\theta(x) + \sum \rho_i(x)}.$$

From (13) we see that $\theta(x) = 0$ only if $x \epsilon G_i$ for some i. But then $\rho_i(x) > 0$. Thus $\theta(x) + \sum \rho_i(x) > 0$ for every $x \epsilon X$ so that g is a continuous mapping. Since Y is convex, $g(X) \subset Y$. From (5) it follows that, when $x \epsilon A$, $\rho_i(x) = 0$ for every i and $\theta(x) > 0$. Thus for $x \epsilon A$, $g(x) = f(x)$; hence $g \epsilon \mathfrak{F}$.

Now we show that $g \epsilon \mathfrak{F}_\epsilon(C_1, C_2)$.

(I) Since A is compact, $g(C_1\cdot A)$ and $g(C_2\cdot A)$ are also compact. Hence $\overline{g(C_1\cdot A)}\cdot\overline{g(C_2\cdot A)} = g(C_1\cdot A)\cdot g(C_2\cdot A)$. Since $g \mid A$ is a homeomorphism and $C_1\cdot C_2 = 0$, it follows that $\overline{g(C_1\cdot A)}\cdot\overline{g(C_2\cdot A)} = 0$.

(II) By (12) and (14) when $x \epsilon X - N_\epsilon$ the point $g(x)$ lies in the interior of the simplex $y_{i_0} \cdots y_{i_u}$, where $G_{i_0}, \cdots, G_{i_u}$ are all the sets G which contain x. Let P_j denote the union of all simplexes $y_{i_0} \cdots y_{i_u}$ such that $C_j\cdot G_{i_0} \cdots G_{i_u} \neq 0$ $(j = 1, 2)$. By (8), P_1 and P_2 are at most n-dimensional, and, by (7), P_1 and P_2 have no vertex in common. Hence, by (10), P_1 and P_2 are disjoint. Since $\overline{g(C_j\cdot(X - N_\epsilon))} \subset \overline{P}_j = P_j$, it follows that $\overline{g(C_1\cdot(X - N_\epsilon))}\cdot\overline{g(C_2\cdot(X - N_\epsilon))} = 0$.

(III) By (11) P_1 and P_2 are disjoint to $f(A) = \overline{g(A)}$. Hence

$$\overline{g(C_1\cdot(X - N_\epsilon))}\cdot\overline{g(C_2\cdot A)} = 0 = \overline{g(C_1\cdot A)}\cdot\overline{g(C_2\cdot(X - N_\epsilon))}.$$

Upon assembling these results we find that $\overline{g(C_1\cdot K_\epsilon)}\cdot\overline{g(C_2\cdot K_\epsilon)} = 0$. Thus $g \epsilon \mathfrak{F}_\epsilon(C_1, C_2)$. Finally we show that $d(f^*, g) < \epsilon$. Clearly

$$g(x) - f^*(x) = \frac{\sum \rho_i(x)(y_i - f^*(x))}{\theta(x) + \sum \rho_i(x)}.$$

By (6) and (9), $d(y_i, f^*(x)) < \epsilon$ whenever $x \epsilon G_i$. But if $x \epsilon X - G_i$, then $\rho_i(x) = 0$. Thus $d(g(x), f^*(x)) < \epsilon$. This completes the proof of (4).

Proof of (1). Let $p = q + n$. Then $p \geqq 2n + 1$, and we need only verify the "span" condition of (4).

Let $y_1, \cdots, y_v$ determine a linear subspace L of R^p of dimension less than n which is disjoint to $f(A)$. If $n = \infty$, then v is finite so that the projection of the span of $f(A)$ and L into the second factor of $R^\infty \times R^\infty$ is contained in a finite linear space. Since this finite linear space has a dense complement in E^∞, it follows that the span of $f(A)$ and L has a dense complement in $Y = E^\infty \times E^\infty$. If n is finite, we need only verify that the dimension of the span of $f(A)$ and L is of dimension less than p. Since the span of $f(A)$ and L is contained in a linear space of dimension $1 + q + \dim L$, this follows inductively from the following: *If B is an at most $(s - 1)$-dimensional subset of an s-dimensional linear subspace L, then the span of B and a point y of R is at most s-dimensional.* If y is in L, the span of B and y is in L and hence has dimension not exceeding $\dim L = s$. If y is not in L, then the relation $\dim \{B \times [0, 1]\} \leqq 1 + \dim B$ and an argument of Hurewicz[10] shows that the dimension of the span $\leqq s$ in this case also.

UNIVERSITY OF ILLINOIS.

[10] W. Hurewicz, *Sur la dimension des produits cartésiens*, Annals of Math., vol. 36(1935), pp. 194–197; especially p. 195.

A GENERALIZATION OF BROUWER'S FIXED POINT THEOREM

By Shizuo Kakutani

The purpose of the present paper is to give a generalization of Brouwer's fixed point theorem (see [1][1]), and to show that this generalized theorem implies the theorems of J. von Neumann ([2], [3]) obtained by him in connection with the theory of games and mathematical economics.

1. The fixed point theorem of Brouwer reads as follows: *if* $x \to \varphi(x)$ *is a continuous point-to-point mapping of an r-dimensional closed simplex S into itself, then there exists an* $x_0 \in S$ *such that* $x_0 = \varphi(x_0)$.

This theorem can be generalized in the following way: Let $\mathfrak{K}(S)$ be the family of all closed convex subsets of S. A point-to-set mapping $x \to \Phi(x) \in \mathfrak{K}(S)$ of S into $\mathfrak{K}(S)$ is called upper semi-continuous if $x_n \to x_0$, $y_n \in \Phi(x_n)$ and $y_n \to y_0$ imply $y_0 \in \Phi(x_0)$. It is easy to see that this condition is equivalent to saying that the graph of $\Phi(x)$: $\sum_{x \in S} x \times \Phi(x)$ is a closed subset of $S \times S$, where $\times$ denotes a Cartesian product. Then the generalized fixed point theorem may be stated as follows:

Theorem 1. *If* $x \to \Phi(x)$ *is an upper semi-continuous point-to-set mapping of an r-dimensional closed simplex S into* $\mathfrak{K}(S)$, *then there exists an* $x_0 \in S$ *such that* $x_0 \in \Phi(x_0)$.

Proof. Let $S^{(n)}$ be the n-th barycentric simplicial subdivision of S. For each vertex x^n of $S^{(n)}$ take an arbitrary point y^n from $\Phi(x^n)$. Then the mapping $x^n \to y^n$ thus defined on all vertices of $S^{(n)}$ will define, if it is extended linearly inside each simplex of $S^{(n)}$, a continuous point-to-point mapping $x \to \varphi_n(x)$ of S into itself. Consequently, by Brouwer's fixed point theorem, there exists an $x_n \in S$ such that $x_n = \varphi_n(x_n)$. If we now take a subsequence $\{x_{n_\nu}\}$ ($\nu = 1, 2, \cdots$) of $\{x_n\}$ ($n = 1, 2, \cdots$) which converges to a point $x_0 \in S$, then this x_0 is a required point.

In order to prove this, let Δ_n be an r-dimensional simplex of $S^{(n)}$ which contains the point x_n. (If x_n lies on the lower-dimensional simplex of $S^{(n)}$, then Δ_n is not uniquely determined. In this case, let Δ_n be any one of these simplexes.) Let $x_0^n, x_1^n, \cdots, x_r^n$ be the vertices of Δ_n. Then it is clear that the sequence $\{x_i^{n_\nu}\}$ ($\nu = 1, 2, \cdots$) converges to x_0 for $i = 0, 1, \cdots, r$, and we have $x_n = \sum_{i=0}^{r} \lambda_i^n x_i^n$ for suitable $\{\lambda_i^n\}$ ($i = 0, 1, \cdots, r$; $n = 1, 2, \cdots$) with $\lambda_i^n \geqq 0$ and $\sum_{i=0}^{r} \lambda_i^n = 1$. Let us further put $y_i^n = \varphi_n(x_i^n)$ ($i = 0, 1, \cdots, r$;

Received January 21, 1941.

[1] Numbers in brackets refer to the list of references at the end.

$n = 1, 2, \cdots$). Then we have $y_i^n \in \Phi(x_i^n)$ and $x_n = \varphi_n(x_n) = \sum_{i=0}^{r} \lambda_i^n y_i^n$ for $n = 1, 2, \cdots$. Let us now take a further subsequence $\{n'_\nu\}$ $(\nu = 1, 2, \cdots)$ of $\{n_\nu\}$ $(\nu = 1, 2, \cdots)$ such that $\{y_i^{n'_\nu}\}$ and $\{\lambda_i^{n'_\nu}\}$ $(\nu = 1, 2, \cdots)$ converge for $i = 0, 1, \cdots, r$, and let us put $\lim_{\nu\to\infty} y_i^{n'_\nu} = y_i^0$ and $\lim_{\nu\to\infty} \lambda_i^{n'_\nu} = \lambda_i^0$ for $i = 0, 1, \cdots, r$. Then we have clearly $\lambda_i^0 \geqq 0$, $\sum_{i=0}^{n} \lambda_i^0 = 1$ and $x_0 = \sum_{i=0}^{r} \lambda_i^0 y_i^0$. Since $x_i^{n'_\nu} \to x_0$, $y_i^{n'_\nu} \in \Phi(x_i^{n'_\nu})$ and $y_i^{n'_\nu} \to y_i^0$ for $i = 0, 1, \cdots, r$, we must have, by the upper semi-continuity of $\Phi(x)$, $y_i^0 \in \Phi(x_0)$ for $i = 0, 1, \cdots, r$, and this implies, by the convexity of $\Phi(x_0)$, that $x_0 = \sum_{i=0}^{r} \lambda_i^0 y_i^0 \in \Phi(x_0)$. Thus the proof of Theorem 1 is completed.

Remark. It is easy to see that Brouwer's fixed point theorem is a special case of Theorem 1 when each $\Phi(x)$ consists only of one point $\varphi(x)$. In this case, the upper semi-continuity of $\Phi(x)$ is nothing but the continuity of $\varphi(x)$.

As an immediate consequence of Theorem 1 we have

COROLLARY. *Theorem 1 is also valid even if S is an arbitrary bounded closed convex set in a Euclidean space.*

Proof. Take a closed simplex S' which contains S as a subset, and consider a continuous retracting point-to-point mapping $x \to \psi(x)$ of S' onto S. ($\psi(x) = x$ for any $x \in S$ and $\psi(x) \in S$ for any $x \in S'$.) Then $x \to \Phi(\psi(x))$ is clearly an upper semi-continuous point-to-set mapping of S' into $\mathfrak{R}(S) \subseteq \mathfrak{R}(S')$. Hence, by Theorem 1, there exists an $x_0 \in S'$ such that $x_0 \in \Phi(\psi(x_0))$. Since $\Phi(\psi(x_0)) \subseteq S$, we must have $x_0 \in S$ and consequently, by the retracting property of $\psi(x)$, $x_0 \in \Phi(x_0) \subseteq S$. This completes the proof of the corollary.

2. THEOREM 2. *Let K and L be two bounded closed convex sets in the Euclidean spaces R^m and R^n respectively, and let us consider their Cartesian product $K \times L$ in R^{m+n}. Let U and V be two closed subsets of $K \times L$ such that for any $x_0 \in K$ the set U_{x_0}, of all $y \in L$ such that $(x_0, y) \in U$, is non-empty, closed and convex, and such that for any $y_0 \in L$ the set V_{y_0}, of all $x \in K$ such that $(x, y_0) \in V$, is non-empty, closed and convex. Under these assumptions, U and V have a common point.*

Proof. Put $S = K \times L$, and let us define a point-to-set mapping $z \to \Phi(z)$ of S into $\mathfrak{R}(S)$ as follows: $\Phi(z) = V_y \times U_x$ if $z = (x, y)$. Since U and V are both closed by assumption, $\Phi(z)$ is clearly upper semi-continuous. Hence, by the corollary of Theorem 1, there exists a point $z_0 \in K \times L$ such that $z_0 \in \Phi(z_0)$. In other words, there exists a pair of points x_0 and y_0, $x_0 \in K$, $y_0 \in L$ such that $(x_0, y_0) \in V_{y_0} \times U_{x_0}$ or equivalently, $x_0 \in V_{y_0}$ and $y_0 \in U_{x_0}$. This means that $z_0 = (x_0, y_0) \in U \cdot V$, and the proof of Theorem 2 is completed.

Remark. Theorem 2 is due to J. von Neumann [3], who proved this by using a notion of integral in Euclidean spaces. The proof given above is simpler.

This theorem has applications to the problems of mathematical economics as was shown by J. von Neumann.

THEOREM 3. *Let $f(x, y)$ be a continuous real-valued function defined for $x \in K$ and $y \in L$, where K and L are arbitrary bounded closed convex sets in two Euclidean spaces R^m and R^n. If for every $x_0 \in K$ and for every real number α, the set of all $y \in L$ such that $f(x_0, y) \leqq \alpha$ is convex, and if for every $y_0 \in L$ and for every real number β, the set of all $x \in K$ such that $f(x, y_0) \geqq \beta$ is convex, then we have*

$$\max_{x \in K} \min_{y \in L} f(x, y) = \min_{y \in L} \max_{x \in K} f(x, y).$$

Proof. Let U and V be the sets of all $z_0 = (x_0, y_0) \in K \times L$ such that $f(x_0, y_0) = \min_{y \in L} f(x_0, y)$ and $f(x_0, y_0) = \max_{x \in K} f(x, y_0)$ respectively. Then it is easy to see that both U and V satisfy the conditions of Theorem 2. Hence, by Theorem 2, there exists a point $z_0 = (x_0, y_0) \in K \times L$ such that $z_0 \in U \cdot V$ or equivalently, $f(x_0, y_0) = \min_{y \in L} f(x_0, y) = \max_{x \in K} f(x, y_0)$. Consequently, we have $\min_{y \in L} \max_{x \in K} f(x, y) \leqq \max_{x \in K} f(x, y_0) = f(x_0, y_0) = \min_{y \in L} f(x_0, y) \leqq \max_{x \in K} \min_{y \in L} f(x, y)$. Since it is clear that we have $\min_{y \in L} \max_{x \in K} f(x, y) \geqq \max_{x \in K} \min_{y \in L} f(x, y)$, the proof of Theorem 3 is completed.

Remark. Theorem 3 is one of the fundamental theorems in the theory of games developed by J. von Neumann [2].

In concluding this paper I should like to express my hearty thanks to Dr. A. D. Wallace for his kind discussions on this problem. He has also obtained analogous results for trees. (A. D. Wallace [4].)

REFERENCES

1. B. KNASTER, C. KURATOWSKI AND S. MAZURKIEWICZ, *Ein Beweis des Fixpunktsatzes für n-dimensionale Simplexe*, Fund. Math., vol. 14(1929), pp. 132–137.
2. J. VON NEUMANN, *Zur Theorie der Gesellschaftsspiele*, Math. Ann., vol. 100(1928), pp. 295–320.
3. J. VON NEUMANN, *Über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes*, Erg. Math. Kolloqu., vol. 8(1937), pp. 73–83.
4. A. D. WALLACE, *A fixed point theorem for trees*, forthcoming in Bull. Amer. Math. Soc., vol. 47(1941).

INSTITUTE FOR ADVANCED STUDY.

THE DOUBLE LAPLACE INTEGRAL

By Dorothy L. Bernstein

The Laplace integral

$$\int_0^\infty e^{-sx} F(x)\, dx,$$

where x is real and s is real or complex, has been the subject of many extensive investigations. More recently, the Laplace-Stieltjes integral

$$\int_0^\infty e^{-sx}\, d\varphi(x)$$

has been studied (see, for example, the series of papers by D. V. Widder which appears in the Transactions of the American Mathematical Society, vols. 31, 33, 36, 39); this integral includes as special cases the ordinary Laplace integral and the Dirichlet series.

The object of this paper is to investigate the analogous integral for functions of two variables:

$$\int_0^\infty \int_0^\infty e^{-sx-ty} d_x d_y \varphi(x, y).$$

In some cases, the results are direct generalizations of the one variable theory; in others, the methods and results are quite different.

I. Introduction

1. Let $\varphi(x)$ be a complex-valued function of a real variable x, defined on the closed interval (a, b). The definitions of a function of bounded variation and of the total variation $V_\varphi[a, b]$, usually given for real-valued functions, apply equally well here. (See, for example, [18], p. 325.) If $\varphi(x) = \theta(x) + i\psi(x)$, where $\theta(x)$ and $\psi(x)$ are real, $\varphi(x)$ is of bounded variation on (a, b) if and only if $\theta(x)$ and $\psi(x)$ are of bounded variation. Hence the properties established for real-valued functions ([18], pp. 325–330) can readily be extended to complex-valued functions.

Although the concepts of positive variation $P_\varphi[a, b]$ and negative variation $N_\varphi[a, b]$ apply only to real-valued functions, the following extension of the fundamental property of functions of bounded variation is valid:

Lemma 1. *A necessary and sufficient condition that $\varphi(x)$ be of bounded variation is that it be expressible as a linear combination, with constant complex coeffi-*

Received July 26, 1939; in revised form, January 22, 1941. This paper is based, in part, upon the author's dissertation written at Brown University under the direction of Professor J. D. Tamarkin.

cients, of a finite number of non-negative, bounded functions, all monotone increasing or all monotone decreasing, on (a, b).[1]

If $\varphi(x)$ is of bounded variation on (a, b), the function $v(x)$ defined thus:

$$v(a) = 0, \qquad v(x) = V_\varphi[a, x] \text{ for } a < x \leqq b,$$

is a non-negative, bounded, monotone increasing function of x on (a, b). We can write, as a particular case of the lemma,

$$\varphi(x) = \varphi(a) + \sum_{j=1}^{4} c_j \varphi_j(x), \tag{1}$$

where the $\varphi_j(x)$ are monotone increasing and

$$| c_j | = 1; \qquad 0 \leqq \varphi_j(x) \leqq v(x) \tag{2}$$

for all x on (a, b).[2]

The classical definition of the Stieltjes integral

$$\int^b f(x)\, d\varphi(x)$$

for a finite interval ([18], p. 538), and its elementary properties are applicable to complex-valued functions, with the above definition of bounded variation.[3]

2. Let $\varphi(x)$ be defined for all $x \geqq 0$ and be a function of bounded variation in every closed interval $(0, X)$. Then $v(x)$ is defined for all $x \geqq 0$, and $\varphi(x)$ can be written in the form (1). But in general, $v(x)$ will not be bounded in $0 \leqq x < \infty$, even though $\varphi(x)$ is bounded; hence the $\varphi_j(x)$ need no longer be bounded.

If $v(x) \leqq M$ for all $x \geqq 0$, then $\varphi(x)$ is called a *function of bounded variation on* $(0, \infty)$. In this case the $\varphi_j(x)$ are also bounded, and from (1) and the fact that $v(x)$ and $\varphi_j(x)$ are monotone, it follows that the limits of $v(x)$, $\varphi_j(x)$, and $\varphi(x)$ all exist as x becomes infinite. Furthermore, the statement of Lemma 1 remains true if (a, b) is replaced by $(0, \infty)$.

If $f(x)$ and $\varphi(x)$ are defined for all $x \geqq 0$, and $f(x)$ is continuous and $\varphi(x)$ is of bounded variation in every $(0, X)$, the function $S(x)$ defined by the equations

$$S(0) = 0, \qquad S(X) = \int_0^X f(x)\, d\varphi(x) \qquad \text{for } X > 0$$

[1] A function of bounded variation might be defined as a function having the property specified in the lemma ([8], p. 9).

[2] Let $\varphi_j(a) = 0$, and $\varphi_1(x) = P_\theta[a, x]$, $\varphi_2(x) = N_\theta[a, x]$, $\varphi_3(x) = P_\psi[a, x]$, $\varphi_4(x) = N_\psi[a, x]$ for $a < x \leqq b$.

[3] Statements of the properties we shall use may be found in [16], pp. 269–272, where this integral is called a "norm" integral. This article contains various other definitions of the Stieltjes integral and the relationships between them.

exists for all $x \geqq 0$ and is of bounded variation in every $(0, X)$ ([20], p. 283). If $\lim_{X\to\infty} S(X)$ exists, we write

$$\lim_{X\to\infty} S(X) = \int_0^\infty f(x)\, d\varphi(x) \tag{3}$$

and say that the infinite integral *converges.*

If

$$\int_0^\infty f(x)\, dv(x) \tag{4}$$

converges, then we say that the integral in (3) *converges absolutely.* Since ([20], p. 283)

$$V_S[a, b] = \int_a^b f(x)\, dv(x), \tag{5}$$

the convergence of (4) is equivalent to the bounded variation of $S(x)$ on $(0, \infty)$.

An absolutely convergent integral also converges, but not conversely.[4] A simple case of an absolutely convergent integral occurs when $\varphi(x)$ is of bounded variation on $(0, \infty)$, and $f(x)$ is bounded on $(0, \infty)$, as well as being continuous in every $(0, X)$; in fact, this is often the only case for which an infinite integral is defined.

3. Let $\varphi(x, y)$ be a complex-valued function defined on the closed rectangle R: $a \leqq x \leqq b$, $c \leqq y \leqq d$.[5] If for every net N consisting of a set of lines of the form

$$x = x_i \ (i = 1, 2, \cdots, m), \qquad a = x_1 < x_2 < \cdots < x_m = b;$$

$$y = y_j \ (j = 1, 2, \cdots, n), \qquad c = y_1 < y_2 < \cdots < y_n = d,$$

the quantity

$$T_\varphi(N) = \sum_{i=1}^{m} \sum_{j=1}^{n} \left| \varphi(x_i, y_j) - \varphi(x_i, y_{j-1}) - \varphi(x_{i-1}, y_j) + \varphi(x_{i-1}, y_{j-1}) \right|$$

is bounded, $\varphi(x, y)$ is said to have *bounded second variation on R.* The least upper bound of $T_\varphi(N)$, for all nets N, is called the *total variation of $\varphi(x, y)$ on R,* and will be denoted by $V_\varphi[a, b; c, d]$. A function which has bounded second variation on R will be said to *belong to class V on R.*

Similarly $\varphi(x, y)$ *belongs to class H on R,* if $\varphi(x, y)$ is a function of class V, if for some x_0 on (a, b), $\varphi(x_0, y)$ is a function of bounded variation in y on (c, d), and if, for some y_0 on (c, d), $\varphi(x, y_0)$ is a function of bounded variation in x on

[4] Compare the situation in the finite case ([16], p. 272).

[5] There are several possible definitions of functions of bounded variations in two variables, see [3] and [4]. The definitions given below are extensions of definitions V and H given in [3], p. 825.

(a, b). If we denote the total x-variation on (a, b) of $\varphi(x, \bar{y})$ by $V'_\varphi[a, b; \bar{y}]$, then for any $\bar{y}$ not equal to y_0

$$V'_\varphi[a, b; \bar{y}] \leqq V'_\varphi[a, b; y_0] + V_\varphi[a, b; c, d], \tag{6}$$

and hence $\varphi(x, y)$ is of bounded variation in x for every y on (c, d). A similar statement is true concerning $V''_\varphi[\bar{x}; c, d]$, the y-variation on (c, d) of $\varphi(\bar{x}, y)$.

If $\varphi(x, y) \in V$ on R, and if $\varphi(x, c) = 0$, $a \leqq x \leqq b$; $\varphi(a, y) = 0$, $c \leqq y \leqq d$, then $\varphi(x, y)$ *belongs to class* V_0 on R. Since $V'_\varphi[a, b; c] = 0$ and $V''_\varphi[a; c, d] = 0$, V_0 is a subclass of H as well as of V.

If $\varphi(x, y) = \theta(x, y) + i\psi(x, y)$, a necessary and sufficient condition that $\varphi(x, y)$ belong to class V (or H or V_0) on R is that $\theta(x, y)$ and $\psi(x, y)$ belong to class V (or H or V_0) on R. Hence the properties established for real-valued functions can be extended to complex-valued functions. (See [3], [4], and, for class H, [26].)

4. If $\varphi(x, y) \in V$ on R, define the variation function $v(x, y)$ thus:

$$v(x, c) = 0,\ a \leqq x \leqq b; \qquad v(a, y) = 0 \text{ for } c \leqq y \leqq d;$$
$$v(x, y) = V_\varphi[a, x; c, y] \qquad \text{for } a < x \leqq b,\ c < y \leqq d.$$

If $\varphi(x, y) \in H$ on R, define two additional functions $v'(x, y)$ and $v''(x, y)$ in the following manner:

$$v'(a, y) = 0, \qquad v'(x, y) = V'_\varphi[a, x; y] \text{ for } a < x \leqq b;$$
$$v''(x, c) = 0, \qquad v''(x, y) = V''_\varphi[a; c, y] \text{ for } c < y \leqq d.$$

From (6), it follows that

$$v'(x, y) \leqq v(x, y) + v'(x, c).$$

Similarly,

$$v''(x, y) \leqq v(x, y) + v''(a, y).$$

Using the notation

$$\Delta_{10}\varphi(x, y) = \varphi(x_2, y) - \varphi(x_1, y),$$
$$\Delta_{01}\varphi(x, y) = \varphi(x, y_2) - \varphi(x, y_1),$$
$$\Delta_{11}\varphi(x, y) = \varphi(x_2, y_2) - \varphi(x_2, y_1) - \varphi(x_1, y_2) + \varphi(x_1, y_1),$$

where $x_2 > x_1$, $y_2 > y_1$, we easily see that

$$|\Delta_{10}\varphi(x, y)| \leqq \Delta_{10}v'(x, y),$$
$$|\Delta_{01}\varphi(x, y)| \leqq \Delta_{01}v''(x, y),$$
$$|\Delta_{11}\varphi(x, y)| \leqq \Delta_{11}v(x, y)$$

everywhere in R, when $\varphi(x, y) \in H$. When $\varphi(x, y) \in V_0$, we have

$$|\varphi(x, y)| \leqq v'(x, y) \leqq v(x, y),$$
$$|\varphi(x, y)| \leqq v''(x, y) \leqq v(x, y)$$

for all (x, y) in R.

The following important property, which corresponds to Lemma 1, can be proved from the analogous property for real-valued functions ([4], p. 718):

LEMMA 2. *A necessary and sufficient condition that $\varphi(x, y)$ belong to class H on R is that it be expressible as a linear combination, with constant complex coefficients, of a finite number of bounded functions $\varphi_j(x, y)$ each satisfying conditions*

$$(\alpha) \qquad \varphi_j(x, y) \geqq 0, \quad \Delta_{10}\varphi_j(x, y) \geqq 0, \quad \Delta_{01}\varphi_j(x, y) \geqq 0, \quad \Delta_{11}\varphi_j(x, y) \geqq 0$$

or each satisfying conditions

$$(\beta) \qquad \varphi_j(x, y) \geqq 0, \quad \Delta_{10}\varphi_j(x, y) \leqq 0, \quad \Delta_{01}\varphi_j(x, y) \leqq 0, \quad \Delta_{11}\varphi_j(x, y) \geqq 0$$

everywhere in R.

Note that when $\varphi(x, y) \in V$ on R, since $V_\varphi[a', b'; c', d']$ is an additive function of the subrectangles of R, it follows that $v(x, y)$ satisfies (α) and is bounded; hence $v(x, y) \in V$ on R.

The particular representation corresponding to (1), when $\varphi(x, y) \in H$ on R, is

$$(7) \qquad \varphi(x, y) = \varphi(a, c) + \sum_{j=1}^{r} c_j \varphi_j(x, y),$$

where the $\varphi_j(x, y)$ satisfy (α) and

$$(8) \qquad |c_j| = 1, \qquad 0 \leqq \varphi_j(x, y) \leqq v(x, y) + v'(x, c) + v''(a, y)$$

everywhere on R.[6]

5. For complex-valued functions of two variables, the Stieltjes integral

$$(9) \qquad \int_a^b \int_c^d f(x, y)\, d_x\, d_y\, \varphi(x, y)$$

is defined exactly as for real-valued functions.[7]

LEMMA 3. *If $\varphi(x, y) \in V$ on R, we can associate with it a function $\psi(x, y)$ with the following properties:*

(1) *$\psi(x, y) \in V_0$ on R.*

(2) *$V_\psi[a, b; c, d] = V_\varphi[a, b; c, d]$.*

(3) *For a given $f(x, y)$, the existence of either of the integrals*

$$\int_a^b \int_c^d f(x, y)\, d_x\, d_y\, \psi(x, y), \qquad \int_a^b \int_c^d f(x, y)\, d_x\, d_y\, \varphi(x, y)$$

implies the existence of the other and the equality of both.

[6] Young ([26], [27], [28]) uses $v(x, y) + v'(x, c) + v''(a, y)$ as his variation function instead of $v(x, y)$.

[7] We shall use the "unrestricted" integral, as defined by Clarkson, [11], p. 930. Besides the references given in this paper, see Young [27], p. 280.

The proof is immediate if we set

$$\psi(x, y) = \varphi(x, y) - \varphi(x, c) - \varphi(a, y) + \varphi(a, c)$$

and use the fact that $\Delta_{11}\psi(x, y) = \Delta_{11}\varphi(x, y)$.[8]

Hence, when dealing with integrals, it is often sufficient to consider functions of class V_0 instead of class V; as we have seen, this will afford a considerable simplification of proofs. Furthermore, this lemma enables us to extend theorems on integrals established for φ belonging to H to the case of φ belonging to V. For example, we can say that (9) exists not only when f is continuous and $\varphi \epsilon V$ ([11], p. 933), but also when $f \epsilon H$ and φ is continuous and $\varphi \epsilon V$.[9]

6. An important subclass of H is formed by functions $f(x, y)$ which are absolutely continuous.[10] Every absolutely continuous function is an "indefinite integral"; that is,

$$f(x, y) = \int_a^x \int_c^y F(u, v)\, du\, dv + \int_a^x G(u)\, du + \int_b^y H(v)\, dv + C$$

for all (x, y) on R, where $F(x, y)$ is summable (Lebesgue) in R, $G(x)$ is summable on (a, b) and $H(y)$ is summable on (c, d). (See [10], p. 654; [18], pp. 592, 615.)

If $\varphi(x, y) \epsilon V_0$ and $f(x, y)$ is absolutely continuous,

$$\int_a^b \int_c^d \varphi(x, y)\, d_x d_y f(x, y) = \int_a^b \int_c^d \varphi(x, y) \frac{\partial^2 f(x, y)}{\partial x\, \partial y}\, dx\, dy.$$

(See [28], p. 33.) If we use this fact and the corresponding result for one variable, the general formula for integration by parts ([27], p. 281; [18], p. 666) becomes in this case

$$\begin{aligned}\int_a^b \int_c^d f(x, y)\, d_x d_y\, \varphi(x, y) &= f(b, d)\varphi(b, d) - \int_a^b \varphi(x, d) \frac{\partial f(x, d)}{\partial x}\, dx \\ &\quad - \int_c^d \varphi(a, y) \frac{\partial f(a, y)}{\partial y}\, dy + \int_a^b \int_c^d \varphi(x,y) \frac{\partial^2 f}{\partial x\, \partial y}\, dx\, dy.\end{aligned} \tag{10}$$

7. Let $\varphi(x, y)$ be defined for all (x, y) in $\bar{R}$: $0 \leqq x < \infty$, $0 \leqq y < \infty$, and let it be a function of class V in every R_{XY} : $0 \leqq x \leqq X$, $0 \leqq y \leqq Y$. In general, the total variation $v(x, y)$ will not be bounded in $\bar{R}$, even though $\varphi(x, y)$ is bounded. If there does exist an M such that $v(x, y) \leqq M$ in $\bar{R}$, then we say that $\varphi(x, y)$ *belongs to class V on* $\bar{R}$.

[8] See [17], pp. 319–320, where functions of class V_0 are called functions of bounded variation.

[9] [27], p. 281 for $\varphi \epsilon H$. The properties of double Stieltjes integrals used in this paper are found in [11], [14], [17], [27], [28].

[10] According to the definitions given (for real-valued functions) by Carathéodory ([10], p. 653). Such a function is absolutely continuous according to Hobson's definition ([18], p. 346) and $f(x, c)$ and $f(a, y)$ are absolutely continuous. Hence ([18], p. 346), $f(x, y) \epsilon H$ on R.

Suppose that $\varphi(x, y) \in H$ in every R_{XY}. Then $v(x, y)$, $v'(x, y)$, $v''(x, y)$ are defined in $\bar{R}$, but are not bounded in general. It is still possible to represent $\varphi(x, y)$ as a linear combination of functions satisfying (α) (for example, (7) is still valid), but the functions will no longer be bounded. And $\varphi(x, y)$ cannot be represented as a linear combination of functions satisfying (β).[11]

If $\varphi(x, y) \in H$ in every R_{XY}, and if there exists an M such that $v(x, y) \leqq M$ for all $x \geqq 0$, $y \geqq 0$, $v'(x, \bar{y}) \leqq M$ for all $x \geqq 0$ and some $\bar{y} \geqq 0$, $v''(\bar{x}, y) \leqq M$ for some $\bar{x} \geqq 0$ and all $y \geqq 0$, then we say that $\varphi(x, y)$ *belongs to H on $\bar{R}$*. One can easily show that $v'(x, y)$ and $v''(x, y)$ are bounded for all $x \geqq 0$, $y \geqq 0$. Since $v(x, y)$ is bounded and satisfies (α), the following limits exist: $\lim_{y\to\infty} v(x, y) = \eta(x)$, $\lim_{x\to\infty} v(x, y) = \zeta(y)$, $\lim_{x,y\to\infty} v(x, y) = \lim_{x\to\infty} \eta(x) = \lim_{y\to\infty} \zeta(y)$, where the double limit is taken as x and y become infinite, simultaneously but independently. From the one variable case, $\lim_{x\to\infty} v'(x, c)$ and $\lim_{y\to\infty} v''(a, y)$ exist.

From (7) and (8) it follows that the analogous limits for $\varphi_i(x, y)$ and $\varphi(x, y)$ all exist, when $\varphi \in H$ on $\bar{R}$; furthermore, Lemma 2 remains true when R is replaced by $\bar{R}$.

If $\varphi(x, y) \in V$ on $\bar{R}$ and $\varphi(x, 0) = \varphi(0, y) = 0$, for all $x \geqq 0$, $y \geqq 0$, we say that $\varphi(x, y)$ *belongs to class V_0 on $\bar{R}$*. As in the finite case, V_0 is a subclass of H.

8. Let $f(x, y)$ and $\varphi(x, y)$ be defined for all $x \geqq 0$, $y \geqq 0$, and let $f(x, y)$ be continuous and $\varphi(x, y) \in H$ in every $R_{XY}: 0 \leqq x \leqq X$, $0 \leqq y \leqq Y$. Set $S(X, 0) = 0$, $S(0, Y) = 0$ and

$$S(X, Y) = \int_0^Y \int_0^X f(x, y)\, d_x d_y \varphi(x, y)$$

for $X > 0$, $Y > 0$. Then $S(x, y)$ is defined for all $x \geqq 0$, $y \geqq 0$, and belongs to class V_0 in every R_{XY}.[12] If $\lim_{X, Y\to\infty} S(X, Y)$ exists, we write

$$\lim_{X, Y\to\infty} S(X, Y) = \int_0^\infty \int_0^\infty f(x, y)\, d_x\, d_y\, \varphi(x, y) \tag{11}$$

and say that the double integral *converges*.[13] The finite integrals $S(X, Y)$ are called the *sections* of the infinite integral (11), which we can denote by $J(f, \varphi)$.

If $J(f, \varphi)$ exists, it need not be true that $| S(X, Y) |$ is bounded in $\bar{R}$. When this is the case, we say that $J(f, \varphi)$ *converges boundedly*. If

$$\int_0^\infty \int_0^\infty | f(x, y) |\, d_x\, d_y\, v(x, y) = J(| f |, v) \tag{12}$$

[11] In $\bar{R}$, a function $w(x, y)$ can satisfy (α) without being bounded, but if it satisfies (β), $0 \leqq w(x, y) \leqq w(0, 0)$ so that it is necessarily bounded. This is an important distinction which is apt to be overlooked since for finite closed rectangles, boundedness follows from either (α) or (β).

[12] The proof follows the corresponding proof in one variable.

[13] For the analogous case of non-absolutely convergent Lebesgue integrals in $\bar{R}$, see [9], pp. 311–318.

converges, then we say that $J(f, \varphi)$ *converges absolutely*. Since

$$V_S[0, 0; X, Y] = \int_0^Y \int_0^X | f(x, y) | \, d_x \, d_y \, v(x, y),$$

a necessary and sufficient condition that (12) exists is that $S(x, y)$ be of bounded variation in $\bar{R}$: $0 \leqq x < \infty, 0 \leqq y < \infty$.

If $J(f, \varphi)$ converges absolutely, it converges boundedly, and

$$J(f, \varphi) = \lim_{X, Y \to \infty} S(X, Y) = \lim_{X \to \infty} \lim_{Y \to \infty} S(X, Y) = \lim_{Y \to \infty} \lim_{X \to \infty} S(X, Y).$$

A simple test for absolute convergence is the following: If $| f(x, y) | \leqq | g(x, y) |$ and $J(g, \varphi)$ converges absolutely, then $J(f, \varphi)$ converges absolutely.

9. If $f(x, y) = g(x) \cdot h(y)$, the following problem arises: when can the double Stieltjes integral be replaced by repeated integrals? For finite rectangles, we have the following result:[14]

THEOREM. *Let $f(x, y) = g(x) \cdot h(y)$, where each of these is continuous. Let $\varphi(x, y) \in V_0$ on R. Then*

$$\theta(x) = \int_c^d h(y) \, d_y \, \varphi(x, y)$$

and

$$\lambda(y) = \int_a^b g(x) \, d_x \, \varphi(x, y)$$

are functions of bounded variation on $a \leqq x \leqq b$ and $c \leqq y \leqq d$ respectively and

$$\int_a^b \int_c^d g(x) \cdot h(y) \, d_x \, d_y \, \varphi(x, y) = \int_a^b g(x) \, d\theta(x) = \int_c^d h(y) \, d\lambda(y).$$

If $f(x, y) = g(x) \cdot h(y)$ is continuous and $\varphi(x, y) \in V_0$ in every R_{XY}, the corresponding theorem does not hold in $\bar{R}$, even in the case of absolute convergence of $J(f, \varphi)$.[15] We shall define a repeated infinite integral as follows.

If

$$\bar{\theta}(x) = \lim_{Y \to \infty} \int_0^Y h(y) \, d_y \, \varphi(x, y)$$

exists and is a function of bounded variation in x in every $[0, X]$, and if

$$\lim_{X \to \infty} \int_0^X g(x) \, d\bar{\theta}(x)$$

[14] The proof for real-valued functions and "restricted" integrals is given in [14], p. 245; in this case the "restricted" and "unrestricted" integrals are equal ([11], p. 931).

[15] Note the difference in the case of the Lebesgue integral.

exists, this limit is called the repeated integral

$$\int_0^\infty g(x)\,d_x\left[\int_0^\infty h(y)\,d_y\,\varphi(x,y)\right]. \tag{13}$$

Similarly, we define

$$\int_0^\infty h(y)\,d_y\left[\int_0^\infty g(x)\,d_x\varphi(x,y)\right] = \lim_{Y\to\infty}\int_0^Y h(y)\,d_y\left[\lim_{X\to\infty}\int_0^X g(x)\,d_x\varphi(x,y)\right]. \tag{14}$$

A useful concept in the theory of double series is that of regular convergence, which means convergence in the Pringsheim sense and convergence of each row and each column. For if a series converges regularly, it converges boundedly and it can be summed by rows or by columns ([19], pp. 49, 51; [2], p. 979). The analogous definition for double integrals is not very useful, since the existence of $J(f, \varphi)$ and of $\bar{\theta}(x)$ does not imply the existence of (13) nor does it imply the boundedness of $S(x, y)$. Hence we shall not use the concept of regular convergence but merely require the existence of (13) or (14) when necessary.

The following theorem illustrates a particularly simple case:

THEOREM. *If $g(x)$ and $h(y)$ are bounded continuous functions of x and y respectively for all $x \geqq 0$, $y \geqq 0$, and if $\varphi(x, y) \in V_0$ on $\bar{R}$, then*

$$\int_0^\infty g(x)\,d_x\varphi(x,y), \qquad \int_0^\infty h(y)\,d_y\,\varphi(x,y),$$

$$\int_0^\infty\int_0^\infty g(x)h(y)\,d_x d_y\,\varphi(x,y)$$

converge absolutely and (13) *and* (14) *exist and are equal to $J(f, \varphi)$.*

II. The double Laplace integral. Preliminary theorems

1. Let $\varphi(x, y) \in V_0$ in every R_{XY}, and let $s = \sigma + i\lambda$ and $t = \tau + i\mu$ be two independent complex variables. Then e^{-sx-ty} is a continuous function of class H in every R_{XY}, and the integral

$$\int_0^Y\int_0^X e^{-sx-ty}\,d_x d_y\,\varphi(x,y)$$

exists for all $X > 0$, $Y > 0$. Set $L(X, Y; s, t)$ equal to the above integral when it exists, and $L = 0$ when either X or Y is 0.

The infinite integral

$$\int_0^\infty\int_0^\infty e^{-sx-ty}\,d_x d_y\,\varphi(x,y)$$

is called the *double Laplace integral*, and will be denoted by $\mathfrak{L}(\varphi; s, t)$. The function which this integral defines, when it exists, will be denoted by $f(s, t)$. The above integral includes as special cases the Lebesgue integral

$$\int_0^\infty\int_0^\infty e^{-sx-ty}F(x,y)\,dx\,dy,$$

where $F(x, y)$ is a function summable in every R_{XY}, and the double Dirichlet series[16]

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{mn} e^{-p_m s - q_n t}.$$

2. We shall consider in this section certain consequences of the following condition:

(A) There exist real numbers M, a_0, b_0 such that for all $x \geqq 0$, $y \geqq 0$

$$| \varphi(x, y) | \leqq M e^{a_0 x + b_0 y}.$$

THEOREM 1. *If* (A) *is true,* $\mathfrak{L}(\varphi; s, t)$ *converges and its sections are bounded for all* (s, t) *such that* $\sigma > a_0$, $\tau > b_0$; *for all such* (s, t),

$$f(s, t) = st \cdot g(s, t),$$

where $g(s, t)$ *is the function defined by the absolutely convergent Lebesgue integral*

$$\int_0^\infty \int_0^\infty e^{-sx-ty} \varphi(x, y)\, dx\, dy. \tag{15}$$

Proof. By the formula for integration by parts, namely, (10),

$$\begin{aligned} L(X, Y; s, t) = \int_0^Y \int_0^X e^{-sx-ty} d_x d_y \varphi(x, y) &= e^{-sX-tY} \varphi(X, Y) \\ + s \int_0^X e^{-sx-tY} \varphi(x, Y)\, dx &+ t \int_0^Y e^{-sX-ty} \varphi(X, y)\, dy \\ &+ st \int_0^Y \int_0^X e^{-sx-ty} \varphi(x, y)\, dx\, dy. \end{aligned} \tag{16}$$

Let $\sigma = a_0 + h$, $\tau = b_0 + k$, where $h > 0$, $k > 0$. Then by (A), for all $x \geqq 0$, $y \geqq 0$,

$$| e^{-sx-ty} \varphi(x, y) | = e^{-\sigma x - \tau y} | \varphi(x, y) | \leqq M e^{-hx-ky}. \tag{17}$$

Hence

$$\begin{aligned} | L(X, Y; s, t) | \leqq M \Big[e^{-hX-kY} &+ e^{-kY} | s | \int_0^X e^{-hx}\, dx \\ &+ e^{-hX} | t | \int_0^Y e^{-ky}\, dy + | st | \int_0^Y \int_0^X e^{-hx-ky}\, dx\, dy \Big]. \end{aligned} \tag{18}$$

[16] Double Dirichlet series have been thoroughly investigated by Kojima [21] (Adams [2] contains references), Leja [22], and others. Some work has been done on double Riemann integrals by Vignaux ([23], [24], and papers referred to in them), and on slightly more general integrals by Biggeri ([5], [6], [7]), who has also considered double Dirichlet series. Stieltjes integrals of real-valued functions have been considered to some extent by Durañona y Vedia and Trejo.

Since $e^{-hX} < 1$, $e^{-kY} < 1$ for all X, Y, it follows from (18) upon performing integrations, that

$$(19) \qquad | L(X, Y; s, t) | \leqq M\left[1 + \frac{|s|}{h} + \frac{|t|}{k} + \frac{|st|}{hk}\right]$$

for all $X \geqq 0$, $Y \geqq 0$. Thus the sections of $\mathfrak{L}(\varphi; s, t)$ are bounded.

As X and Y become infinite, the first three terms in the last member of (18) tend to 0, and the integral in the last term converges absolutely, by the comparison test. Hence the rest of the theorem is true.

THEOREM 2.[17] *If* (A) *is true,* $\mathfrak{L}(\varphi; s, t)$ *converges uniformly with respect to* s *and* t, *and has its sections uniformly bounded in every region*

$$R: \quad \sigma \geqq a_0 + \delta, \qquad |s| \leqq H\sigma; \qquad \tau \geqq b_0 + \eta, \qquad |t| \leqq K\eta,$$

where δ *and* η *are arbitrarily small positive numbers and* H, K *are finite.*

Proof. We need to prove that given $\epsilon > 0$, there exist X_0 and Y_0, depending only upon ϵ, such that

$$| L(X_2, Y_2; s, t) - L(X_1, Y_1; s, t) | < \epsilon$$

for all $X_2 > X_1 \geqq X_0$, $Y_2 > Y_1 \geqq Y_0$, (s, t) in R.

In the notation of Theorem 1, $e^{-hX_2} < e^{-hX_1} < 1$, $e^{-kY_2} < e^{-kY_1} < 1$. Applying formula (16) at (X_1, Y_1) and at (X_2, Y_2), using (17), and performing integrations, we obtain

$$| L(X_2, Y_2; s, t) - L(X_1, Y_1; s, t) |$$
$$< 2M\left[e^{-hX_1-kY_1} + \frac{|s|}{h}e^{-kY_1} + \frac{|t|}{k}e^{-hX_1} + \frac{|st|}{hk}(e^{-hX_1} + e^{-kY_1})\right].$$

Since $h \geqq \delta$ and $k \geqq \eta$ in R,

$$\frac{|s|}{h} \leqq H\frac{\sigma}{h} = H\left(1 + \frac{a_0}{h}\right) \leqq H\left(1 + \frac{|a_0|}{\delta}\right) = \bar{H},$$
$$\frac{|t|}{k} \leqq K\left(1 + \frac{|b_0|}{\eta}\right) = \bar{K}.$$

For all (s, t) in R,

$$| L(X_2, Y_2; s, t) - L(X_1, Y_1; s, t) |$$
$$< 2M[e^{-\delta X_1-\eta Y_1} + \bar{H}e^{-\eta Y_1} + \bar{K}e^{-\delta X_1} + \bar{H}\bar{K}e^{-\delta X_1} + \bar{H}\bar{K}e^{-\eta Y_1}];$$

this last quantity does not depend upon s and t, and it can be made less than ϵ, by taking X_1 and Y_1 sufficiently large.

From (19), for all (s, t) in R,

$$| L(X, Y; s, t) | \leqq M[1 + \bar{H} + \bar{K} + \bar{H}\bar{K}].$$

[17] Biggeri ([5], p. 811) establishes uniform convergence in a more general type of region, but under the assumption that the real parts of s and t are always equal.

COROLLARY. *If* (A) *is true,* $\mathfrak{L}(\varphi; s, t)$ *converges uniformly with respect to* s *and* t, *and has its sections uniformly bounded in every region*

$$\bar{R}:\quad \sigma \geqq a_0 + \delta, \qquad -\alpha < \lambda < \alpha; \qquad \tau \geqq b_0 + \eta, \qquad -\beta < \mu < \beta,$$

where δ *and* η *are arbitrarily small positive numbers, and* α *and* β *are finite positive numbers.*

The proof is immediate, since $\bar{R}$ is contained in R_1 for sufficiently large H and K.

If (A) is true, it can be shown exactly as in the one variable case ([25], pp. 699–703) that for every t such that $\tau > b_0$ and for each $x \geqq 0$, the integral

$$\int_0^\infty e^{-ty}\, d_y\varphi(x, y)$$

exists, its sections are bounded by $M_1 e^{a_0 x}$, and the function $k(x; t)$ which it defines is equal to $t \cdot l(x; t)$, where $l(x; t)$ is defined by the absolutely convergent integral

$$\int_0^\infty e^{-ty}\varphi(x, y)\, dy.$$

Since (15) may be written as a repeated integral for $\sigma > a_0$, $\tau > b_0$,

$$(20)\quad f(s, t) = st\int_0^\infty \int_0^\infty e^{-sx-ty}\varphi(x, y)\, dx\, dy$$

$$= st\int_0^\infty e^{-sx}l(x; t)\, dt = s\int_0^\infty e^{-sx}k(x; t)\, dx.$$

However, $k(x; t)$ is not, in general, of bounded variation in x in $(0, X)$, so that $f(s, t)$ cannot be written as a repeated Stieltjes integral.

Similar results are true concerning $\bar{k}(y; s)$, defined by

$$\int_0^\infty e^{-sx}\, d_x\varphi(x, y).$$

3. If we define functions $v(x, y)$, $v'(x, y)$, $v''(x, y)$ as equal to $V_\varphi[0, 0; x, y]$, $V'_\varphi[0, x; y]$, $V''_\varphi[x; 0, y]$, respectively for $x > 0$, $y > 0$, and equal to 0 when x or y is 0, then

$$(21)\quad |\varphi(x, y)| \leqq v'(x, y) \leqq v(x, y), \qquad |\varphi(x, y)| \leqq v''(x, y) \leqq v(x, y)$$

for all $x \geqq 0$, $y \geqq 0$. In this section we shall consider the consequences of replacing condition (A) by similar conditions regarding v', v'', v.

THEOREM 3. *If there exist* M, a_0, b_0 *such that*

$$(22)\qquad v'(x, y) \leqq Me^{a_0x+b_0y}$$

for all $x \geqq 0$, $y \geqq 0$, the results of §2 *are valid; furthermore, for every $\sigma > a_0$, $\tau > b_0$, $k(x; t)$ is of bounded x-variation in every $(0, X)$ and*

$$f(s, t) = \int_0^\infty e^{-sx} d_x k(x; t). \tag{23}$$

Proof. From (21), (A) is a result of (22), so that the first part of the theorem is true.

Since for $\tau > b_0$ and any x in $(0, X)$

$$\int_0^\infty e^{-\tau y} \mid \varphi(x, y) \mid dy$$

converges,

$$\sum_{i=1}^n \mid l(x_i, t) - l(x_{i-1}, t) \mid \leqq \int_0^\infty e^{-\tau y} \cdot \sum_{i=1}^n \mid \varphi(x_i, y) - \varphi(x_{i-1}, y) \mid dy$$

$$\leqq \int_0^\infty e^{-b_0 y} v'(x, y)\, dy \leqq Me^{a_0 X} \cdot \int_0^\infty e^{(b_0 - \tau) y} dy = \frac{Me^{a_0 X}}{\tau - b_0}$$

for every net of the form $0 = x_0 < x_1 < \cdots < x_m = X$. Hence $l(x; t)$ and $k(x; t) = t \cdot l(x; t)$ are of bounded variation in $(0, X)$.

$$\mid k(x; t) \mid \leqq V_k[0, x; t] \leqq \frac{M \mid t \mid}{\tau - b_0} e^{a_0 x}$$

for all $x \geqq 0$. As in one variable, then ([25], pp. 699–703),

$$\int_0^\infty e^{-sx} d_x\, k(x; t)$$

exists for $\sigma > a_0$, $\tau > b_0$, and is equal to

$$s \int_0^\infty e^{-sx} k(x; t)\, dx.$$

Substituting in (20), we get (23).

THEOREM 4. *If there exist M, a_0, b_0 such that*

$$v''(x, y) \leqq Me^{a_0 x + b_0 y} \tag{24}$$

for all $x \geqq 0$, $y \geqq 0$, the results of §2 *are valid; for every $\sigma > a_0$, $\tau > b_0$, $\bar{k}(y; s)$ is of bounded y-variation in every $(0, Y)$ and*

$$f(s, t) = \int_0^\infty e^{-ty} d_y\, \bar{k}(y; s).$$

THEOREM 5. *If there exist M, a_0, b_0 such that*

$$v(x, y) \leqq Me^{a_0 x + b_0 y} \tag{25}$$

for all $x \geqq 0$, $y \geqq 0$, $\mathfrak{L}(\varphi; s, t)$ converges absolutely for all (s, t) with $\sigma > a_0$, $\tau > b_0$. The results of Theorems 3 and 4 are valid and the integrals defining $k(x; t)$ and $\bar{k}(y; s)$ converge absolutely.

Proof. Since $v(x, y) \in V_0$, by Theorem 1 $\mathfrak{L}(v; s, t)$ converges for all (s, t) such that $\sigma > a_0$, $\tau > b_0$. Taking real values of s and t, we see that

$$\mathfrak{L}(v; \sigma, \tau) = \int_0^\infty \int_0^\infty e^{-\sigma x - \tau y} d_x d_y v(x, y)$$

converges for $\sigma > a_0$, $\tau > b_0$; that is, $\mathfrak{L}(\varphi; s, t)$ converges absolutely.

From (21), (22) and (24) are consequences of (25), so that the results of Theorems 3 and 4 hold. From (24), it easily follows that

$$\int_0^\infty e^{-\tau y} d_y v''(x, y)$$

converges for $\tau \geqq b_0$ and every $x \geqq 0$; that is,

$$k(x; t) = \int_0^\infty e^{-ty} d_y \varphi(x, y)$$

converges absolutely. Similarly, from (22), the integral defining $\bar{k}(y; s)$ converges absolutely.

Remark. The theorems of §§2 and 3 are true if the hypotheses hold for all $x \geqq 0$, $y \geqq 0$, except for a finite subregion D.

III. Regions of convergence of the double Laplace integral

1. THEOREM 6.[18] *Let the sections of*

$$\mathfrak{L}(\varphi; s, t) = \int_0^\infty \int_0^\infty e^{-sx-ty} d_x d_y \varphi(x, y)$$

be bounded at a finite point (s_0, t_0); *that is,*

$$(26) \qquad | L(X, Y; s_0, t_0) | = \left| \int_0^Y \int_0^X e^{-s_0 x - t_0 y} d_x d_y \varphi(x, y) \right| \leqq M$$

for all $X \geqq 0$, $Y \geqq 0$. *Then* $\mathfrak{L}(\varphi; s, t)$ *converges boundedly for every* (s, t) *in* $S: \sigma > \sigma_0$, $\tau > \tau_0$.

Proof. Let $s = s_0 + m$, $t = t_0 + n$, where $R(m) = h > 0$, $R(n) = k > 0$. Then ([28], p. 33)

$$L(X, Y; s, t) = \int_0^Y \int_0^X e^{-mx-ny} \cdot e^{-s_0 x - t_0 y} d_x d_y \varphi(x, y)$$
$$= \int_0^Y \int_0^X e^{-mx-ny} d_x d_y L(x, y; s_0, t_0).$$

[18] Theorems analogous to Theorem 6 and its corollary are given in [1], p. 407, [24], p. 291, and stated in [12], p. 319.

Since $\varphi(x, y) \epsilon V_0$, $L(x, y; s_0, t_0) \epsilon V_0$ in every R_{XY}, and $| L(x, y; s_0, t_0) | < M$ for all $x \geqq 0$, $y \geqq 0$. Hence proceed as in Theorem 1, using

$$(27) \qquad | e^{-mx-ny} L(x, y; s_0, t_0) | \leqq Me^{-hx-ky}$$

instead of (17).

COROLLARY. *If $\mathfrak{L}(\varphi; s, t)$ converges boundedly at (s_0, t_0), it converges boundedly at every (s, t) such that $\sigma > \sigma_0$, $\tau > \tau_0$.*

2. THEOREM 7.[19] *Let the sections of $\mathfrak{L}(\varphi; s, t)$ be bounded at (s_0, t_0). Then $\mathfrak{L}(\varphi; s, t)$ converges uniformly and has its sections uniformly bounded in every region*

$$S_1: \sigma \geqq \sigma_0 + \delta, | s - s_0 | \leqq H(\sigma - \sigma_0); \qquad \tau \geqq \tau_0 + \eta, | t - t_0 | \leqq K(\tau - \tau_0),$$

where δ, η are arbitrarily small and positive, and H and K are finite.

If we use (27), the proof is analogous to that of Theorem 2; in S_1

$$h \geqq \delta, k \geqq \eta, \qquad \frac{| m |}{h} = \frac{| s - s_0 |}{\sigma - \sigma_0} < H, \qquad \frac{| n |}{k} \leqq K.$$

Given any finite region D which is interior to S, H, K, δ, η can be selected such that D is interior to S_1, and hence $\mathfrak{L}(\varphi; s, t)$ converges uniformly in D. Thus we have the following theorem:

THEOREM 8. *If the sections of $\mathfrak{L}(\varphi; s_0, t_0)$ are bounded, $\mathfrak{L}(\varphi; s, t)$ represents an analytic function $f(s, t)$ within the domain S: $\sigma > \sigma_0$, $\tau > \tau_0$, and in this region,*

$$\frac{\partial^{h+k} f(s, t)}{\partial s^h \partial t^k} = \int_0^\infty \int_0^\infty e^{-sx-ty}(-x)^h(-y)^k \, d_x \, d_y \, \varphi(x, y).$$

The proof is given in [12], page 322. Another proof can easily be constructed along the lines of the proof given by Widder in [25] for the one variable case, based on the extension of Vitali's theorem given by Kojima, [21], page 380.

3. THEOREM 9. *If the sections of $\mathfrak{L}(\varphi; \sigma_1, \tau_1)$ are bounded where (σ_1, τ_1) is a real finite point, there exists a constant M such that, for all $x \geqq 0$, $y \geqq 0$,*

$$(28) \qquad | \varphi(x, y) | \leqq Me^{\sigma_2 x + \tau_2 y},$$

where $\sigma_2 = \max(0, \sigma_1)$, $\tau_2 = \max(0, \tau_1)$.

Proof.

$$\varphi(X, Y) = \int_0^Y \int_0^X e^{\sigma_1 x + \tau_1 y} \cdot e^{-\sigma_1 x - \tau_1 y} \, d_x \, d_y \, \varphi(x, y)$$

$$= \int_0^Y \int_0^X e^{\sigma_1 x + \tau_1 y} \cdot d_x \, d_y \, L(x, y; \sigma_1, \tau_1).$$

[19] Analogous results are given in [1], p. 407, [5], p. 317, and stated in [12], p. 321.

Integrating by parts, and using the hypothesis that $| L(x, y; \sigma_1, \tau_1) | < M$, we get

$$
(29)\qquad \begin{aligned} | \varphi(X, Y) | \leqq M \Big[e^{\sigma_1 X + \tau_1 Y} + e^{\tau_1 Y} | \sigma_1 | \int_0^X e^{\sigma_1 x}\, dx \\ + e^{\sigma_1 X} | \tau_1 | \int_0^Y e^{\tau_1 y}\, dy + | \sigma_1 \tau_1 | \int_0^Y \int_0^X e^{\sigma_1 x + \tau_1 y}\, dx\, dy \Big]. \end{aligned}
$$

If $\sigma_1 \geqq 0$,

$$
| \sigma_1 | \int_0^X e^{\sigma_1 x}\, dx = e^{\sigma_1 X} - 1 < e^{\sigma_1 X}.
$$

If $\sigma_1 < 0$,

$$
| \sigma_1 | \int_0^X e^{\sigma_1 x}\, dx = 1 - e^{\sigma_1 X} < 1.
$$

For all real σ_1,

$$
| \sigma_1 | \int_0^X e^{\sigma_1 x}\, dx < \max (1, e^{\sigma_1 X}) = e^{\sigma_2 X}.
$$

Similarly,

$$
| \tau_1 | \int_0^Y e^{\tau_1 y}\, dy < e^{\tau_2 X}.
$$

The result then follows from (29).

For any point $(\bar{s}, \bar{t})$ where $\bar{\sigma} \geqq 0$, $\bar{\tau} \geqq 0$, and where

$$
| L(X, Y; \bar{s}, \bar{t}) | < M,
$$

$$
| \varphi(x, y) | < M e^{\bar{\sigma} x + \bar{\tau} y}.
$$

Therefore the representations of §2, part II, hold for all $\sigma > \bar{\sigma}$, $\tau > \bar{\tau}$. But if $\bar{\sigma} < 0$ or $\bar{\tau} < 0$, these representations will hold true only for those $\sigma > \bar{\sigma}$, $\tau > \bar{\tau}$ which have $\sigma > 0$, $\tau > 0$.

4. From the corollary to Theorem 6, it follows that if $\mathfrak{L}(\varphi; s, t)$ does not converge boundedly at (s_1, t_1), it does not converge boundedly at any (s, t) such that $\sigma < \sigma_1$, $\tau < \tau_1$. Hence there exist two real numbers c and k with the following properties:

B_1 : $\mathfrak{L}(\varphi; s, t)$ converges boundedly for $R(s) > c$, $R(t) > k$.
B_2 : $\mathfrak{L}(\varphi; s, t)$ does not converge boundedly for $R(s) < c$, $R(t) < k$.

The numbers c and k are known as associated abscissas of bounded convergence, since if they are both finite, the region of convergence of $\mathfrak{L}(\varphi; s, t)$ consists of two half-planes, one in the s-plane to the right of the line $\sigma = c$ and the other

in the t-plane to the right of the line $\tau = k$.[20] The numbers c and k are not unique, but neither are they independent of each other. If (c_1, k_1) and (c_2, k_2) are two pairs of abscissas of bounded convergence and $c_1 < c_2$, it follows from B_1 and B_2 that $k_1 \geqq k_2$. Given $\varphi(x, y)$, the determination of formulas for c and k has been discussed in some detail; we shall give some of the more important results.

5. THEOREM 10.[21] *A necessary and sufficient condition that the positive numbers c and k be associated abscissas of bounded convergence is that*

$$\overline{\lim_{x+y\to\infty}} \frac{\log |\varphi(x, y)|}{cx + ky} = 1. \tag{30}$$

Proof. Let (c, k) be a pair of positive constants satisfying (30). Let $(\bar{s}, \bar{t})$ be a point such that $\bar{\sigma} > c$, $\bar{\tau} > k$. Then there exists $\epsilon > 0$ such that $\bar{\sigma} > c(1 + \epsilon)$, $\bar{\tau} > k(1 + \epsilon)$; given ϵ, there exists finite $T > 0$ such that

$$\frac{\log |\varphi(x, y)|}{cx + ky} < 1 + \epsilon,$$

$$|\varphi(x, y)| < e^{(1+\epsilon)(cx+ky)}$$

for all $x + y > T$. Since $0 \leqq x + y \leqq T$ is a closed finite subregion, by the remark on page 473, we can apply Theorem 1 and obtain bounded convergence of $\mathfrak{L}(\varphi; s, t)$ at all (s, t) such that $\sigma > c(1 + \epsilon)$, $\tau > k(1 + \epsilon)$ and hence at $(\bar{s}, \bar{t})$. This establishes property B_1.

Let (s', t') be a point at which $\sigma' < c$, $\tau' < k$, and suppose $\mathfrak{L}(\varphi; s, t)$ converges boundedly at (s', t'). It is sufficient to consider the case $\sigma' > 0$, $\tau' > 0$. Choose ϵ such that $\sigma' < \sigma' + \epsilon < \sigma' + 2\epsilon < c$, $\tau' < \tau' + \epsilon < \tau' + 2\epsilon < k$. Then by Theorem 6, $\mathfrak{L}(\varphi; s, t)$ converges boundedly at $(\sigma' + \epsilon, \tau' + \epsilon)$, and by Theorem 9:

$$|\varphi(x, y)| < Ke^{(\sigma'+\epsilon)x+(\tau'+\epsilon)y} \tag{31}$$

for all $x \geqq 0$, $y \geqq 0$. Let δ be such that $\sigma' + 2\epsilon < c(1 - \delta)$, $\tau' + 2\epsilon < k(1 - \delta)$. From (30), for certain $x + y$ as large as desired,

$$\frac{\log |\varphi(x, y)|}{cx + ky} > 1 - \delta,$$

$$|\varphi(x, y)| > e^{(\sigma'+2\epsilon)x+(\tau'+2\epsilon)y}$$

and this contradicts (31) for $x + y$ sufficiently large. Hence $\mathfrak{L}(\varphi; s, t)$ does not converge boundedly at (s', t'). Thus B_2 is established.

[20] If $\mathfrak{L}(\varphi; s, t)$ converges boundedly for every (s, t), we can take $c = +\infty$, $k = +\infty$. If $\mathfrak{L}(\varphi; s, t)$ converges boundedly for no (s, t), we can take $c = -\infty$, $k = -\infty$. For the cases where c or k is infinite, see the parallel discussion by Kojima in [21].

[21] Given in [12], p. 323; the corresponding result for Dirichlet series is in [21], p. 150.

To show the necessity of the condition, let $c > 0$, $k > 0$ be a pair of abscissas of bounded convergence, and set

$$r = \overline{\lim_{x+y\to\infty}} \frac{\log |\varphi(x, y)|}{cx + ky}.$$

If $r > 0$,

$$\overline{\lim_{x+y\to\infty}} \frac{\log |\varphi(x, y)|}{cx + ky} = 1;$$

and by the first part of the proof (cr, kr) are a pair of abscissas of bounded convergence. This is impossible if $r \neq 1$, since either $c < cr$, $k < kr$, or $cr < c$, $kr < k$. If $r \leqq 0$, for $0 < \epsilon < 1$, there exists T such that

$$\frac{\log |\varphi(x, y)|}{cx + ky} < r + \epsilon \leqq \epsilon, \qquad |\varphi(x, y)| < e^{xc\epsilon+yk\epsilon}$$

for $x + y > T$. By Theorem 1 then, the integral converges boundedly for $\sigma > c\epsilon$, $\tau > k\epsilon$. This is impossible by property B_2. Hence $r = 1$ is the only possible value.

Corollary.[22] *A necessary and sufficient condition that $\mathfrak{L}(\varphi; s, t)$ converge boundedly for $\sigma > c$, $\tau > k$, where $c > 0$, $k > 0$, is that*

$$\overline{\lim_{x+y\to\infty}} \frac{\log |\varphi(x, y)|}{cx + ky} \leqq 1;$$

Theorem 11.[23] *A necessary and sufficient condition that real numbers* **c** *and* **k** *be associated abscissas of bounded convergence is that*

$$\overline{\lim_{x+y\to\infty}} \frac{\log \left| \int_0^y \int_0^x e^{(1-c)u+(1-k)v}\, d_u\, d_v\, \varphi(u, v) \right|}{x + y} = 1. \tag{32}$$

Proof. Let $s = S + m$, $t = T + n$ where m and n are real, so that $\Re(S) = \sigma - m$, $\Re(T) = \tau - n$. Let

$$G(X, Y) = \int_0^Y \int_0^X e^{-mx-ny}\, d_x\, d_y\, \varphi(x, y).$$

Then

$$\int_0^Y \int_0^X e^{-sx-ty}\, d_x\, d_y\, \varphi(x, y) = \int_0^Y \int_0^X e^{-Sx-Ty}\, d_x\, d_y\, G(x, y).$$

Hence if $\mathfrak{L}(\varphi; s, t)$ converges boundedly at a certain point $(\bar{s}, \bar{t})$, $\mathfrak{L}(G; S, T)$ converges boundedly at the corresponding point $\bar{S} = \bar{s} + m$, $\bar{T} = \bar{t} + n$, and conversely. Thus if $\mathfrak{L}(G; S, T)$ has a pair of abscissas of bounded convergence

[22] [12], p. 325.
[23] Stated in [12], p. 326.

(c', k'), $\mathfrak{L}(\varphi; s, t)$ has $(c' - m, k' - n)$ as a pair of abscissas of bounded convergence.

Let $m = c - 1$, $n = k - 1$. Suppose c and k satisfy (32). Then

$$\overline{\lim_{x+y\to\infty}} \frac{\log | G(x, y) |}{x + y} = 1 \tag{33}$$

and by the previous theorem, $\mathfrak{L}(G; S, T)$ has (1, 1) as a pair of convergence abscissas. Hence $\mathfrak{L}(\varphi; s, t)$ has (c, k) as a pair of abscissas of bounded convergence.

Conversely, let (c, k) be a pair of abscissas of bounded convergence of $\mathfrak{L}(\varphi; s, t)$. Then $(c + m = 1, k + n = 1)$ is a pair of bounded convergence abscissas of $\mathfrak{L}(G; S, T)$ and by the previous theorem (33) is true.

6. The following sufficient conditions that (c, k) be abscissas of bounded convergence are useful in finding a pair of abscissas of bounded convergence for a given $\varphi(x, y)$.

THEOREM 12. *Let (m, n) be any real numbers, and set*

$$\overline{\lim_{x+y\to\infty}} \frac{\log \left| \int_0^y \int_0^x e^{-mu-nv}\, d_u\, d_v\, \varphi(u, v) \right|}{x + y} = A. \tag{34}$$

If $A > 0$, then $\mathfrak{L}(\varphi; s, t)$ has abscissas of bounded convergence $(c = A + m$, $k = A + n)$.

Proof. If we use the notation of Theorem 11, it follows from (34), since $A > 0$, that

$$\overline{\lim_{x+y\to\infty}} \frac{\log | G(x, y) |}{Ax + Ay} = 1.$$

Hence $\mathfrak{L}(G; S, T)$ has a pair of bounded convergence abscissas (A, A). Since $\sigma = \Re(S) + m$, $\tau = \Re(T) + n$, $\mathfrak{L}(\varphi; s, t)$ has a pair of convergence abscissas $(A + m, A + n)$.

THEOREM 13. *Let*

$$K(c) = \overline{\lim_{x+y\to\infty}} \frac{- cx + \log | \varphi(x, y) |}{y}. \tag{35}$$

If $c > 0$, $K(c) > 0$, then $(c, K(c))$ form a pair of abscissas of bounded convergence of $\mathfrak{L}(\varphi; s, t)$.

Proof. From (35),

$$\overline{\lim_{x+y\to\infty}} \frac{\log | \varphi(x, y) |}{cx + K(c)y} = 1,$$

where $c > 0$, $K(c) > 0$. Hence apply Theorem 10.

It follows from the corollary to Theorem 10, or it can be shown directly, that if c and k are given by any of the formulas of this section, $\mathfrak{L}(\varphi; s, t)$ converges boundedly for $\sigma > c$, $\tau > k$, whether or not c or k is positive.

COROLLARY. *Let $c > 0$ and let n be such that*

$$\overline{\lim_{x+y\to\infty}} \frac{\log e^{-cx} \left| \int_0^y e^{-nv} \, d_v \, \varphi(x, v) \right|}{y} = B > 0.$$

Then $\mathfrak{L}(\varphi; s, t)$ has $\sigma = c$, $\tau = B + n$ as abscissas of bounded convergence.

Proof. In the transformation used in Theorem 12, we take $m = 0$,

$$G(x, y) = \int_0^y e^{-nv} \, d_v \, \varphi(x, v).$$

7.[24] Let (E) represent the totality of all σ such that for every σ of (E), there exists a real τ such that $\mathfrak{L}(\varphi; s, t)$ has bounded sections at some (s, t) with $\Re(s) = \sigma$, $\Re(t) = \tau$. If σ_1 belongs to (E), by Theorem 6, all $\sigma > \sigma_1$ belong to (E), so that if r is the lower bound of these values of σ (r may equal $-\infty$), E consists of the half-line $r < \sigma < \infty$. For every σ of E, consider all τ such that $\mathfrak{L}(\varphi; s, t)$ has bounded sections with $\Re(s) = \sigma$, $\Re(t) = \tau$; let their lower bound be $\eta(\sigma)$.

THEOREM 14. *The curve $\tau = \eta(\sigma)$ is defined for all σ of (E): $r < \sigma < \infty$, and has the following properties:*

(1) *For $\sigma_1 < \sigma_2$, $\eta(\sigma_1) \geqq \eta(\sigma_2)$.*

(2) *$\eta(\alpha\sigma_1 + \beta\sigma_2) \leqq \alpha \cdot \eta(\sigma_1) + \beta \cdot \eta(\sigma_2)$ where $\alpha \geqq 0$, $\beta \geqq 0$, $\alpha + \beta = 1$.*

(3) *If $\eta(\sigma_1) = \eta(\sigma_2)$, when $\sigma_1 < \sigma_2$, then $\eta(\sigma)$ is constant on $\sigma_1 \leqq \sigma < \infty$.*

(4) *$\eta(\sigma)$ is continuous at every point of (E).*

Proof. Property (1) follows from Theorem 6. The lemma which forms the basis of the proof of (2) is as follows: If $\mathfrak{L}(\varphi; s, t)$ is bounded at two finite points (s_1, t_1) and (s_2, t_2), it converges at every (s, t) such that $\sigma > \alpha\sigma_1 + \beta\sigma_2$, $\tau > \alpha\tau_1 + \beta\tau_2$, where $\alpha \geqq 0$, $\beta \geqq 0$, $\alpha + \beta = 1$. The proof is omitted since it is like the corresponding proof for Dirichlet series. Properties (3) and (4) follow from (1) and (2).

Hence the domain

$$D: r < \sigma < \infty, \qquad \tau > \eta(\sigma)$$

is convex. Furthermore, for all (s, t) such that (σ, τ) lies in D, $\mathfrak{L}(\varphi; s, t)$ converges boundedly. For all (s, t) such that (σ, τ) lies in the complementary domain, $\mathfrak{L}(\varphi; s, t)$ does not converge boundedly.

8. Since the absolute convergence of $\mathfrak{L}(\varphi; s, t)$ means the bounded convergence of another Laplace integral $\mathfrak{L}(v; \sigma, \tau)$, the discussion of the region of absolute

[24] This section parallels the discussion given by Leja, [22], pp. 150–152.

convergence is parallel to that given for bounded convergence. The first theorem below, however, is slightly stronger than its parallel.

THEOREM 15. *If $\mathfrak{L}(\varphi; s, t)$ converges absolutely at (s_0, t_0), then $\mathfrak{L}(\varphi; s, t)$ converges absolutely for all (s, t) in $\bar{S}$: $\sigma \geqq \sigma_0$, $\tau > \tau_0$. Furthermore, both $\mathfrak{L}(v; \sigma, \tau)$ and $\mathfrak{L}(\varphi; s, t)$ converge uniformly and have their sections uniformly bounded in $\bar{S}$.*

Proof. For (s, t) in $\bar{S}$,

$$| e^{-sx-ty} | = e^{-\sigma x - \tau y} \leqq e^{-\sigma_0 x - \tau_0 y}.$$

Hence by the comparison test, $\mathfrak{L}(v; \sigma, \tau)$ converges in $\bar{S}$. Let

$$\bar{L}(X, Y; \sigma, \tau) = \int_0^Y \int_0^X e^{-\sigma x - \tau y} \, d_x \, d_y \, v(x, y).$$

If $X_2 > X_1$, $Y_2 > Y_1$,

$$
\begin{aligned}
| L(X_2, Y_2; s, t) &- L(X_1, Y_1; s, t) | \\
&= \left| \int_0^{Y_1} \int_{X_1}^{X_2} e^{-sx-ty} \, d_x \, d_y \, \varphi(x, y) + \int_{Y_1}^{Y_2} \int_0^{X_2} e^{-sx-ty} \, d_x \, d_y \, \varphi(x, y) \right| \\
&\leqq \int_0^{Y_1} \int_{X_1}^{X_2} e^{-\sigma x - \tau y} \, d_x \, d_y \, v(x, y) + \int_{Y_1}^{Y_2} \int_0^{X_2} e^{-\sigma x - \tau y} \, d_x \, d_y \, v(x, y) \\
&= \bar{L}(X_2, Y_2; \sigma, \tau) - \bar{L}(X_1, Y_1; \sigma, \tau) \\
&\leqq \bar{L}(X_2, Y_2; \sigma_0, \tau_0) - L(X_1, Y_1; \sigma_0, \tau_0).
\end{aligned}
$$

The last quantity is independent of (s, t) in $\bar{S}$ and can be made less than any ϵ by taking X_1 and Y_1 sufficiently large. Hence $\mathfrak{L}(\varphi; s, t)$ and $\mathfrak{L}(v; \sigma, \tau)$ converge uniformly for (s, t) in $\bar{S}$.

$$| L(X, Y; s, t) | \leqq \bar{L}(X, Y; \sigma, \tau) \leqq \bar{L}(X, Y; \sigma_0, \tau_0) \leqq \mathfrak{L}(v; \sigma_0, \tau_0) = A$$

for all (s, t) in $\bar{S}$ and all $X \geqq 0$, $Y \geqq 0$. Hence the sections of $\mathfrak{L}(v; \sigma, \tau)$ and $\mathfrak{L}(\varphi; s, t)$ are uniformly bounded.

COROLLARY. *If $\varphi(x, y) \in V_0$ in $\bar{R}$: $0 \leqq x < \infty$, $0 \leqq y < \infty$, $\mathfrak{L}(\varphi; s, t)$ converges absolutely for $\sigma \geqq 0$, $\tau \geqq 0$.*

Using Theorem 9, we can prove

THEOREM 16. *If $\mathfrak{L}(\varphi; s, t)$ converges absolutely at (σ_1, τ_1), then, for all $x \geqq 0$, $y \geqq 0$,*

$$v(x, y) < K e^{\sigma_2 x + \tau_2 y},$$

$$\sigma_2 = \max(0, \sigma_1), \qquad \tau_2 = \max(0, \tau_1).$$

Hence for positive values inside the region of absolute convergence, the representations of Theorem 5 hold.

The definition of abscissas of absolute convergence, and the determination of formulas are analogous to those for bounded convergence. For instance, corresponding to Theorem 10, we have

THEOREM 17. *A necessary and sufficient condition that the positive numbers d and m be associated abscissas of absolute convergence of* $\mathfrak{L}(\varphi; s, t)$ *is that*

$$\overline{\lim_{x+y\to\infty}} \frac{\log v(x, y)}{dx + my} = 1.$$

The region of absolute convergence may be described just as in §7.

9. In the introduction, we have seen that the concept of regular convergence, as defined for series, is inapplicable to integrals; hence the theorems concerning half-planes of regular convergence of double Dirichlet series cannot be generalized. Even if one extends concept of regular convergence to mean the existence of repeated integrals, it is not true that if repeated integrals exist at (s_0, t_0), they will exist for $\sigma > \sigma_0$, $\tau > \tau_0$. One could, of course, obtain half-planes in which $\mathfrak{L}(\varphi; s, t)$ converges and satisfies certain auxiliary conditions, which are analogous to the properties of regularly convergent series. For example, if

$$\int_0^\infty e^{-sx} \, d_x \, k(x; t) = \int_0^\infty e^{-sx} \, d_x \int_0^\infty e^{-ty} \, d_y \, \varphi(x, y) \tag{36}$$

exists and

$$V_k'[0, x; t] \leqq Me^{\sigma x} \tag{37}$$

at (s_0, t_0), then (36) exists and (37) holds for all (s, t) such that $\sigma > \sigma_0$, $\tau > \tau_0$. Using the other repeated integral, also, we could get a generalization of the type discussed. Or abandoning the idea of repeated integrals, one could say that $\mathfrak{L}(\varphi; s, t)$ converges regularly boundedly if $\mathfrak{L}(\varphi; s, t)$ exists and if

$$\int_0^Y \int_0^X e^{-sx-ty} \, d_x \, d_y \, \varphi(x, y), \quad \int_0^Y e^{-sX-ty} \, d_y \, \varphi(X, y), \quad \int_0^X e^{-sx-tY} \, d_x \, \varphi(x, Y)$$

are all bounded. Then one could get half-planes of regular bounded convergence. But these concepts are not of much value, since the auxiliary conditions are so restrictive. It is much simpler to state the conditions explicitly needed to obtain desired theorems. See, for example, Theorem 20.

One could also define uniform bounded convergence (uniform convergence and uniformly bounded sections, with respect to s and t), and show the existence of half-planes of uniform bounded convergence. We shall not give the details here, but shall mention one fundamental result:

THEOREM 18. *If the sections of* $\mathfrak{L}(\varphi; s, t)$ *are uniformly bounded in* λ *and* μ *on* $\sigma = a_0$, $\tau = b_0$, *then* $\mathfrak{L}(\varphi; s, t)$ *converges uniformly and has its sections uniformly bounded, with respect to* λ *and* μ, *on every* $\sigma = a$, $\tau = b$, *where* $a > a_0$, $b > b_0$.

Proof. Let $s_0 = a_0 + i\lambda$, $t_0 = b_0 + i\mu$; $s = s_0 + h$, $t = t_0 + k$, where $h > 0$, $k > 0$.

$$L(X, Y; s, t) = \int_0^Y \int_0^X e^{-hx-ky} \cdot d_x d_y L(X, Y; s_0, t_0) = e^{-hX-kY} L(X, Y; s_0, t_0)$$
$$+ h \int_0^X e^{-hx-kY} L(x, Y; s_0, t_0)\, dx + k \int_0^Y e^{-ky-hX} L(X, y; s_0, t_0)\, dy$$
$$+ hk \int_0^Y \int_0^X e^{-hx-ky} \cdot L(x, y; s_0, t_0)\, dx\, dy.$$

Since $h = a - a_0$, $k = b - b_0$, and since by hypothesis $| L(X, Y; s_0, t_0) | < M$ for all λ and μ, the result follows.

IV. Inversion of the double Laplace integral

1. Let $\varphi(x, y)$ be a function of class V_0 in every R_{XY}. At every (w, z), where $w > 0$, $z > 0$, the double limits $\varphi(w + 0, z + 0)$, $\varphi(w + 0, z - 0)$, $\varphi(w - 0, z + 0)$, $\varphi(w - 0, z - 0)$ exist; since $\varphi(x, y)$ is of bounded variation in each variable separately, the limits $\varphi(w, z + 0)$, $\varphi(w, z - 0)$, $\varphi(w + 0, z)$, $\varphi(w - 0, z)$ exist.[25] If (w, z) is a point of continuity, the above quantities are all equal to $\varphi(w, z)$.

Define a function $\bar{\varphi}_{\rho_1,\rho_2}(w, z)$ as follows, for all $w > 0$, $z > 0$:[26]

$$\bar{\varphi}_{0,0}(w, z) = \tfrac{1}{4}[\varphi(w + 0, z + 0) + \varphi(w + 0, z - 0) + \varphi(w - 0, z + 0) + \varphi(w - 0, z - 0)];$$

$$\bar{\varphi}_{\rho_1,0}(w, z) = \frac{1}{\Gamma(\rho_1)} \int_0^w (w - x)^{\rho_1 - 1} \cdot \tfrac{1}{2}[\varphi(x, z + 0) + \varphi(x, z - 0)]\, dx \qquad (\rho_1 > 0);$$

$$\bar{\varphi}_{0,\rho_2}(w, z) = \frac{1}{\Gamma(\rho_2)} \int_0^z (z - y)^{\rho_2 - 1} \cdot \tfrac{1}{2}[\varphi(w + 0, y) + \varphi(w - 0, y)]\, dy \qquad (\rho_2 > 0);$$

$$\bar{\varphi}_{\rho_1,\rho_2}(w, z) = \frac{1}{\Gamma(\rho_1)\Gamma(\rho_2)} \int_0^z \int_0^w (w - x)^{\rho_1 - 1}(z - y)^{\rho_2 - 1} \varphi(x, y)\, dx\, dy \qquad (\rho_1, \rho_2 > 0).$$

By applying the formula for integration by parts, one can show that

$$\frac{1}{\Gamma(\rho_1 + 1) \cdot \Gamma(\rho_2 + 1)} \int_0^z \int_0^w (w - x)^{\rho_1}(z - y)^{\rho_2}\, d_x d_y\, \varphi(x, y) \tag{38}$$

reduces to $\bar{\varphi}_{\rho_1,\rho_2}(w, z)$, when $\rho_1 > 0$, $\rho_2 > 0$; if (w, z) is a point of continuity of $\varphi(x, y)$, (38) becomes $\bar{\varphi}_{\rho_1,\rho_2}(w, z)$ when $\rho_1 = 0$ or when $\rho_2 = 0$.

In the following sections we shall obtain expressions for $\bar{\varphi}_{\rho_1,\rho_2}(w, z)$ in terms of the function $f(s, t)$ defined by the Laplace integral $\mathfrak{L}(\varphi; s, t)$.[27] When double

[25] By [28], p. 31, and Lemma 2 of the introduction.

[26] $\bar{\varphi}_{\rho_1,\rho_2}(w, z)$ is a generalization of a fractional derivative, when $\rho_i > 0$. (See [25], p. 710.) In particular, for integral values of ρ_i, it gives the successive repeated integrals of $\varphi(x, y)$.

[27] For inversion formulas in one variable, see [25], pp. 708–711.

infinite integrals are used, the Cauchy principal value will be understood. That is,

$$\int_{b-i\infty}^{b+i\infty}\int_{a-i\infty}^{a+i\infty} h(s, t)\, ds\, dt \qquad \text{means} \qquad \lim_{\alpha,\beta\to\infty} \int_{b-i\beta}^{b+i\beta}\int_{a-i\alpha}^{a+i\alpha} h(s, t)\, ds\, dt$$

as α and β become infinite simultaneously but independently.

As in the introduction, $V_\varphi[a, b; c, d]$ indicates the total variation of $\varphi(x, y)$ over a rectangle $a \leqq x \leqq b; c \leqq y \leqq d$, $V'_\varphi[a, b; \bar{y}]$ indicates the variation with respect to x of φ on $a \leqq x \leqq b$, for fixed $\bar{y}$, and $V''_\varphi[\bar{x}; c, d]$ indicates the y-variation of φ on $c \leqq y \leqq d$, for fixed $\bar{x}$.

2. Lemma. *For any $\rho_2 \geqq 0$, and $z - \eta > 0$,*

$$(39) \qquad \left| \int_{b-i\beta}^{b+i\beta} \frac{e^{tz}}{t^{\rho_2}}\, dt \int_{z-\eta}^{z+\eta} e^{-ty}\varphi(x, y)\, dy \right| < P\{V''_\varphi[x; z-\eta, z+\eta] + |\varphi(x, z-\eta)|\},$$

where P does not depend on β or x.

Proof. If $\rho_2 = 0$, the integral in (39) becomes

$$\int_{z-\eta}^{z+\eta} \varphi(x, y)\left[\int_{b-i\beta}^{b+i\beta} e^{t(z-y)}\, dt\right] dy = 2i \int_{z-\eta}^{z+\eta} e^{b(z-y)} \frac{\sin\beta(z-y)}{z-y}\varphi(x, y)\, dy.$$

For all p, q such that $z - \eta \leqq p < q \leqq z + \eta$,

$$\left| \int_p^q e^{b(z-y)} \frac{\sin\beta(z-y)}{z-y}\, dy \right| \leqq 2\pi e^{|b|\eta}.$$

Hence by the theorem of the mean for one variable,[28]

$$\left| 2i\int_{-\eta}^{z+\eta} e^{b(z-y)} \frac{\sin\beta(z-y)}{z-y}\varphi(x, y)\, dy \right| \leqq 2\pi e^{|b|\eta}\{V''_\varphi[x; z-\eta, z+\eta] + |\varphi(x, z-\eta)|\}.$$

If $\rho_2 > 0$, the integral in (39) can be written as

$$(40) \qquad \int_{b-i\beta}^{b+i\beta} \frac{e^{tz}}{t^{\rho_2+1}}\left[\int_{z-\eta}^{z+\eta} te^{-ty}\varphi(x, y)\, dy\right] dt.$$

Since for $z - \eta \leqq p < q \leqq z + \eta$

$$\left| \int_{z-\eta}^{z+\eta} te^{-ty}\, dy \right| \leqq e^{-bz+|b|\eta},$$

by the theorem of the mean,

$$\left| \int_{z-\eta}^{z+\eta} te^{-ty}\varphi(x, y)\, dy \right| \leqq e^{-bz+|b|\eta}\{V''_\varphi[x; z-\eta, z+\eta] + |\varphi(x, z-\eta)|\}.$$

[28] Given in [18], p. 263 for real-valued functions. For complex-valued functions, the M which occurs in the inequality need only be replaced by $4M$.

Hence the absolute value of (40) is less than

$$\left[\int_{-\beta}^{\beta} \frac{e^{bz}\, d\mu}{|b + i\mu|^{\rho_2+1}}\right] \cdot e^{-bz+|b|\eta}\{V''_\varphi[x; z - \eta, z + \eta] + |\varphi(x, z - \eta)|\};$$

since

$$\int_{-\infty}^{\infty} \frac{e^{bz}\, d\mu}{|b + i\mu|^{\rho_2+1}}$$

is finite, for $\rho_2 > 0$, the lemma is established in this case also.

3. THEOREM 19. *Let M, a_0, b_0 be real constants such that*

(A) $$|\varphi(x, y)| < Me^{a_0x+b_0y} \text{ for all } x \geqq 0,\ y \geqq 0.$$

Then for every $\rho_1 \geqq 0$, $\rho_2 \geqq 0$, the formula

$$\bar{\varphi}_{\rho_1,\rho_2}(w, z) = \frac{-1}{4\pi^2}\int_{b-i\infty}^{b+i\infty}\int_{a-i\infty}^{a+i\infty} \frac{e^{sw+tz}f(s, t)}{s^{\rho_1+1}t^{\rho_2+1}}\, ds\, dt \qquad [a > a_0, a \neq 0, b > b_0, b \neq 0]$$

is valid at each point (w, z) such that

(B) $V'_\varphi[w - \delta, w + \delta; y] < Me^{b_0y}$ *for all $y \geqq 0$, and some $\delta > 0$,*

(C) $V''_\varphi[x; z - \eta, z + \eta] < Me^{a_0x}$ *for all $x \geqq 0$, and some $\eta > 0$.*

Proof. Set

$$I(\alpha, \beta) = \frac{-1}{4\pi^2}\int_{b-i\beta}^{b+i\beta}\int_{a-i\alpha}^{a+i\alpha} \frac{e^{sw+tz}f(s, t)}{s^{\rho_1+1}t^{\rho_2+1}}\, ds\, dt.$$

By the corollary to Theorem 2, $f(s, t) = \lim_{X,Y\to\infty} L(X, Y; s, t)$ converges uniformly with respect to s and t in $\sigma = a$, $\tau = b$, $-\alpha < \lambda < \alpha$, $-\beta < \mu < \beta$. Hence $I(\alpha, \beta) = \lim_{X,Y\to\infty} J(X, Y; \alpha, \beta)$, where

$$J(X, Y; \alpha, \beta) = \frac{-1}{4\pi^2}\int_{b-i\beta}^{b+i\beta}\int_{a-i\alpha}^{a+i\alpha} \frac{e^{sw+tz}L(X, Y; s, t)}{s^{\rho_1+1}t^{\rho_2+1}}\, ds\, dt.$$

The theorem will be proved if we can show that the two following statements are true:

(I) $J(X, Y; \alpha, \beta)$ converges to $I(\alpha, \beta)$ uniformly in α and β, for all $\alpha \geqq 0$, $\beta \geqq 0$.

(II) If (M, N) are fixed values of X, Y, where $M \geqq w + 1$, $N \geqq z + 1$,

$$\lim_{\alpha,\beta\to\infty} J(M, N; \alpha, \beta) = \bar{\varphi}_{\rho_1,\rho_2}(w, z).$$

Applying the formula for integration by parts to $L(X, Y; s, t)$ and setting

$$H_\gamma(x, \alpha) = \int_{a-i\alpha}^{a+i\alpha} \frac{e^{s(w-x)}}{s^\gamma}\, ds, \qquad K_\gamma(y, \beta) = \int_{b-i\beta}^{b+i\beta} \frac{e^{t(z-y)}}{t^\gamma}\, dt,$$

we obtain

$$
\begin{aligned}
-4\pi^2 \cdot J(X, Y; \alpha, \beta) &= \varphi(X, Y) H_{\rho_1+1}(X, \alpha) K_{\rho_2+1}(Y, \beta) \\
&\quad + H_{\rho_1+1}(X, \alpha) \int_{b-i\beta}^{b+i\beta} \frac{e^{tz}}{t^{\rho_2}} \left[\int_0^Y e^{-ty} \varphi(X, y)\, dy \right] dt \\
&\quad + K_{\rho_2+1}(Y, \beta) \int_{a-i\alpha}^{a+i\alpha} \frac{e^{sw}}{s^{\rho_1}} \left[\int_0^X e^{-sx} \varphi(x, Y)\, dx \right] ds \\
&\quad + \int_{b-i\beta}^{b+i\beta} \int_{a-i\alpha}^{a+i\alpha} \frac{e^{sw+tz}}{s^{\rho_1} t^{\rho_2}} \left[\int_0^Y \int_0^X e^{-sx-ty} \varphi(x, y)\, dx\, dy \right] ds\, dt.
\end{aligned} \tag{41}
$$

To establish (I), we need to show that given ϵ, there exist X_0, Y_0, depending only on ϵ, such that $| J(X_2, Y_2; \alpha, \beta) - J(X_1, Y_1; \alpha, \beta) | < \epsilon$ for all α and β, and all $X_2 > X_1 \geqq X_0$, $Y_2 > Y_1 \geqq Y_0$.

$$
\begin{aligned}
-4\pi^2 &[J(X_2, Y_2; \alpha, \beta) - J(X_1, Y_2; \alpha, \beta)] \\
&= K_{\rho_2+1}(Y_2, \beta)[H_{\rho_1+1}(X_2, \alpha) \cdot \varphi(X_2, Y_2) - H_{\rho_1+1}(X_1, \alpha) \cdot \varphi(X_1, Y_2)] \\
&\quad + K_{\rho_2+1}(Y_2, \beta) \cdot \int_{X_1}^{X_2} H_{\rho_1}(x, \alpha) \cdot \varphi(x, Y_2)\, dx \\
&\quad + H_{\rho_1+1}(X_2, \alpha) \cdot \int_{b-i\beta}^{b+i\beta} \frac{e^{tz}}{t^{\rho_2}} \left[\int_0^{Y_2} e^{-ty} \cdot \varphi(X_2, y)\, dy \right] dt \\
&\quad + H_{\rho_1+1}(X_1, \alpha) \cdot \int_{b-i\beta}^{b+i\beta} \frac{e^{tz}}{t^{\rho_2}} \left[\int_0^{Y_2} e^{-ty} \cdot \varphi(X_1, y)\, dy \right] dt \\
&\quad + \int_{b-i\beta}^{b+i\beta} \int_{a-i\alpha}^{a+i\alpha} \frac{e^{sw+tz}}{s^{\rho_1} t^{\rho_2}} \left[\int_{X_1}^{X_2} \int_0^{Y_2} e^{-sx-ty} \cdot \varphi(x, y)\, dy\, dx \right] ds\, dt.
\end{aligned} \tag{42}
$$

Since for $| k | > 0$, $| a | > 0$, $\gamma \geqq 0$, $\alpha \geqq 0$,

$$
\left| \int_{a-i\alpha}^{a+i\alpha} \frac{e^{sk}}{s^\gamma}\, ds \right| \leqq \frac{Re^{ak}}{| k |},
$$

where R does not depend on k or α, we have, when $x \geqq X_1 \geqq w + 1$, $y \geqq Y_1 \geqq z + 1$,

$$
H_{\rho_1+1}(x, \alpha) \leqq Ne^{a(w-x)}, \qquad K_{\rho_2+1}(y, \beta) \leqq Ne^{b(z-y)}. \tag{43}
$$

Using these inequalities and (A) of the hypothesis, we see that the first two terms in the right member of (42) can be made arbitrarily small by taking X_1 and Y_1 sufficiently large.

In the third term, the integral is equal to

$$
\begin{aligned}
\int_0^{z-\eta} K_{\rho_2}(y, \beta) \cdot \varphi(X_2, y)\, dy &+ \int_{b-i\beta}^{b+i\beta} \frac{e^{tz}}{t^{\rho_2}} \left[\int_{z-\eta}^{z+\eta} e^{-ty} \cdot \varphi(X_2, y)\, dy \right] dt \\
&+ \int_{z+\eta}^{Y_2} K_{\rho_2}(y, \beta) \cdot \varphi(X_2, y)\, dy.
\end{aligned} \tag{44}
$$

From hypothesis (A), and from the fact that for $0 < y < z - \eta$,

$$| K_{\rho_2}(y, \beta) | \leqq \frac{Ne^{b(z-y)}}{z - y} \leqq \frac{Ne^{b(z-y)}}{\eta},$$

it follows that the absolute value of the first term of (44) is less than $M_1 e^{a_0 X_2}$. From the lemma, the absolute value of the second term is less than

$$P\{ V''_\varphi[X_2 ; z - \eta, z + \eta] + | \varphi(X_2 , z - \eta) | \}.$$

Using hypotheses (A) and (C), we see that this is less than $M_2 e^{a_0 X_2}$. From (A) and from (43), the absolute value of the third term of (44) is less than $M_3 e^{a_0 X_2}$. Hence the absolute value of the third term of (42) is less than

$$Ne^{a(w-X_2)} \cdot \bar{M} e^{a_0 X_2} = Pe^{(a_0 - a) X_2},$$

where P does not depend upon x, y, α, or β. By taking X_2 sufficiently large, this term can be made arbitrarily small.

Similarly, the fourth term of (42) is bounded by $\bar{P} e^{(a_0 - a) X_1}$ and can be made as small as desired by taking X_1 large enough.

The last term can be written as

$$\int_{X_1}^{X_2} H_{\rho_1}(x, \alpha) \cdot \left\{ \int_{b-i\beta}^{b+i\beta} \frac{e^{tz}}{t^{\rho_2}} \left[\int_0^{Y_2} e^{-ty} \varphi(x, y) \, dy \right] dt \right\} dx.$$

The integral in brackets is similar to the integrals discussed, hence the absolute value of this term is less than

$$\int_{X_1}^{X_2} Ne^{a(w-x)} \cdot \bar{M} e^{a_0 x} \, dx = P_1 [e^{(a_0 - a) X_1} - e^{(a_0 - a) X_2}].$$

This quantity can also be made arbitrarily small by taking X_1 and X_2 sufficiently large.

Hence given ϵ, there exist $X_0 \geqq w + 1$ and $Y_0 \geqq z + 1$, depending only upon ϵ, such that

$$| J(X_2 , Y_2 ; \alpha, \beta) - J(X_1 , Y_2 ; \alpha, \beta) | < \tfrac{1}{2}\epsilon$$

for $X_2 > X_1 \geqq X_0$, $Y_2 > Y_0$, and all α and β. By a parallel proof, using (B) instead of (C), and the corresponding lemma,

$$| J(X_1 , Y_2 ; \alpha, \beta) - J(X_1 , Y_1 ; \alpha, \beta) | < \tfrac{1}{2}\epsilon$$

for $X_1 > X'_0$, $Y_2 > Y_1 \geqq Y'_0$, and all α and β, where X'_0 and Y'_0 depend only upon ϵ. These two statements establish (I).

Let M and N be fixed numbers greater than $w + 1$ and $z + 1$ respectively. As in the one variable case, the proof of (II) is different for $\rho_i = 0$ and for $\rho_i > 0$.

Case 1. $\rho_1 > 0$, $\rho_2 > 0$. Changing the order of integration ([28], p. 36), we get

$$J(M, N; \alpha, \beta) = \frac{-1}{4\pi^2} \int_0^N \int_0^M H_{\rho_1+1}(x, \alpha) \cdot K_{\rho_2+1}(y, \beta) \cdot d_x \, d_y \, \varphi(x, y).$$

$$\lim_{\alpha \to \infty} H_{\rho_1+1}(x, \alpha) = \int_{a-i\infty}^{a+i\infty} \frac{e^{s(w-x)}}{s^{\rho_1+1}} \, ds$$

converges uniformly for x in $(0, M)$; its value is $\dfrac{2\pi i(w - x)}{\Gamma(\rho_1 + 1)}$, if $0 < x < w$, and 0 if $x > w$. (See [25], p. 711.) A similar statement holds for

$$\lim_{\beta\to\infty} K_{\rho_2+1}(y, \beta),$$

where y is in $(0, N)$. Hence ([14], p. 453)

$$\begin{aligned}\lim J(M, N; \alpha, \beta) &= \int_0^N \int_0^M \lim_{\alpha\to\infty} H_{\rho_1+1}(x, \alpha) \cdot \lim_{\beta\to\infty} K_{\rho_2+1}(y, \beta) \cdot d_x\, d_y\, \varphi(x, y) \\ &= \frac{1}{\Gamma(\rho_1 + 1)\Gamma(\rho_2 + 1)} \int_0^z \int_0^w (w - x)^{\rho_1}(z - y)^{\rho_2}\, d_x\, d_y\, \varphi(x, y).\end{aligned}$$

And this is (38) and hence $\bar{\varphi}_{\rho_1,\rho_2}(w, z)$.

Case 2. $\rho_1 > 0$, $\rho_2 = 0$. If

$$\psi(y, s) = \int_0^M e^{-sx}\, d_x\, \varphi(x, y),$$

$$\begin{aligned}L(M, N; s, t) &= \int_0^N \int_0^M e^{-sx-ty}\, d_x\, d_y\, \varphi(x, y) = \int_0^N e^{-ty}\, d_y \psi(y, s) \\ &= e^{-tN}\psi(N, s) + t \int_0^N e^{-ty}\psi(y, s)\, dy.\end{aligned}$$

Hence

$$\begin{aligned}J(M, N; \alpha, \beta) &= \left[\frac{1}{2\pi i} \int_{b-i\beta}^{b+i\beta} \frac{e^{t(z-N)}}{t}\, dt\right] \cdot \left[\frac{1}{2\pi i} \int_{a-i\alpha}^{a+i\alpha} \frac{e^{sw}\psi(N, s)}{s^{\rho_1+1}}\, ds\right] \\ &\quad - \frac{1}{4\pi^2} \int_{b-i\beta}^{b+i\beta} \int_{a-i\alpha}^{a+i\alpha} \frac{e^{sw+tz}}{s^{\rho_1+1}} \left[\int_0^N e^{-ty}\psi(y, s)\, dy\right] ds\, dt.\end{aligned}$$

In the first term, α and β are in distinct factors. The limit of the first factor as $\alpha \to \infty$ is 0, since $N \geqq z + 1$; the limit of the second factor as $\beta \to \infty$ is ([25], p. 711)

$$\frac{1}{\Gamma(\rho_1 + 1)} \int_0^w (w - x)^{\rho_1}\, d_x\, \varphi(x, N).$$

Hence the entire term tends to 0 as α and β become infinite. The second term may be written as

$$R(\alpha, \beta) = \frac{1}{2\pi i} \int_{a-i\alpha}^{a+i\alpha} \frac{e^{sw}}{s^{\rho_1+1}}\, W(s, \beta)\, ds,$$

where

$$W(s, \beta) = \int_0^N e^{b(z-y)} \frac{\sin \beta(z - y)}{z - y}\, \psi(y, s)\, dy.$$

If we set

$$\lambda(x, \beta) = \frac{1}{2\pi i}\int_0^N e^{b(z-y)}\,\frac{\sin\beta(z-y)}{z-y}\,\varphi(x, y)\,dy,$$

we can also write

$$W(s, \beta) = \int_0^M e^{-sx}\,d_x\lambda(x, \beta).$$

For every β ([25], p. 708),

$$J(\beta) = \lim_{\alpha\to\infty} R(\alpha, \beta) = \frac{1}{\Gamma(\rho_1 + 1)}\int_0^w (w - x)^{\rho_1}\,d_x\lambda(x, \beta).$$

By the theorem of the mean,

$$|\, W(s, \beta)\,| \leqq V'_\psi[0, N; s]\cdot 2\pi\cdot \bar{M} < P\cdot V_\varphi[0, M; 0, N],$$

where P does not depend upon α or β. Thus

$$|\, R(\alpha, \beta)\,| \leqq \frac{P}{2\pi}\,V_\varphi[0, M; 0, N]\cdot\int_{-\alpha}^{\alpha}\frac{e^{aw}\,d\lambda}{|\,a + i\lambda\,|^{\rho_1+1}},$$

and since

$$\int_{-\infty}^{\infty}\frac{e^{aw}\,d\lambda}{|\,a + i\lambda\,|^{\rho_1+1}}$$

converges, $R(\alpha, \beta)$ converges as $\alpha \to \infty$, uniformly in β, for all $\beta \geqq 0$. Hence given ϵ, there exists α_0 such that

$$(45) \qquad |\, R(\alpha, \beta) - J(\beta)\,| < \tfrac{1}{2}\epsilon$$

for $\alpha > \alpha_0$ and all β.

Now from the case of one variable ([25], p. 708),

$$\lim_{\beta\to\infty} \lambda(x, \beta) = \tfrac{1}{2}[\varphi(x, z + 0) + \varphi(x, z - 0)].$$

And $\lambda(x, \beta)$ can easily be shown to be of bounded variation in x on $(0, w)$, uniformly in β for all $\beta \geqq 0$. Hence

$$\lim_{\beta\to\infty} J(\beta) = \frac{1}{\Gamma(\rho_1 + 1)}\int_0^w (w - x)^{\rho_1}\cdot d_x\tfrac{1}{2}[\varphi(x, z + 0) + \varphi(x, z - 0)].$$

This last quantity reduces to $\bar{\varphi}_{\rho_1,0}(w, z)$ upon integration by parts. Hence given ϵ, there exists β_0 such that for $\beta > \beta_0$

$$(46) \qquad |\, J(\beta) - \bar{\varphi}_{\rho_1,0}(w, z)\,| < \tfrac{1}{2}\epsilon.$$

(45) and (46) establish the fact that

$$\lim_{\alpha,\beta\to\infty} R(\alpha, \beta) = \bar{\varphi}_{\rho_1,0}(w, z).$$

Case 3. $\rho_1 = 0, \rho_2 > 0$. The proof is similar to that of Case 2.

Case 4. $\rho_1 = \rho_2 = 0$. Formula (16) for integration by parts becomes

$$J(M, N; \alpha, \beta) = \frac{-1}{4\pi^2} \varphi(X, Y) \cdot K_1(N, \beta) \cdot H_1(M, \alpha) + H_1(M, \alpha) \cdot \lambda(M, \beta)$$

$$\text{(47)} \qquad + K_1(N, \beta) \cdot \bar{\lambda}(N, \alpha) + \frac{1}{\pi^2} \int_0^N \int_0^M e^{a(w-x)+b(z-y)} \varphi(x, y)$$

$$\cdot \frac{\sin \alpha(w - x)}{w - x} \frac{\sin \beta(z - y)}{z - y} \, dx \, dy,$$

where $\lambda(x, \beta)$ is defined on p. 488 and

$$\bar{\lambda}(y, \beta) = \frac{1}{2\pi i} \int_0^M e^{a(w-x)} \frac{\sin \alpha(w - x)}{w - x} \varphi(x, y) \, dx.$$

Since $M \geqq w + 1, N \geqq z + 1$,

$$\lim_{\alpha \to \infty} H_1(M, \alpha) = 0, \qquad \lim_{\beta \to \infty} K_1(N, \beta) = 0,$$

$$\lim_{\beta \to \infty} \lambda(M, \beta) = \tfrac{1}{2}[\varphi(M, z + 0) + \varphi(M, z - 0)],$$

$$\lim_{\alpha \to \infty} \bar{\lambda}(N, \alpha) = \tfrac{1}{2}[\varphi(w + 0, N) + \varphi(w - 0, N)].$$

Hence the first three terms of the right member of (47) tend to 0 as $\alpha, \beta \to \infty$.

The last term of (47) is a double Dirichlet integral, and it tends to $\bar{\varphi}_{0,0}(w, z)$ as $\alpha, \beta \to \infty$. To prove this, write the term as the sum of four integrals taken over rectangles $R_1[0, w; 0, z]$, $R_2[w, M; 0, z]$, $R_3[0, w; z, N]$, $R_4[w, M; z, N]$, respectively. If we set $\nu = w - x, \theta = z - y$, the first integral becomes

$$J_1(\alpha, \beta) = \frac{1}{\pi^2} \int_0^w \int_0^z e^{a\nu+b\theta} \varphi(w - \nu, z - \theta) \frac{\sin \alpha\nu}{\nu} \frac{\sin \beta\theta}{\theta} \, d\nu \, d\theta.$$

Since $e^{a\nu+b\theta}\varphi(w - \nu, z - \theta)$ is the product of functions of class H, it is also of class H ([4], p. 719). Thus

$$e^{a\nu+b\theta} \varphi(w - \nu, z - \theta) = \sum_{i=1}^{r} c_i \omega_i(\nu, \theta)$$

where for $i = 1, 2, \cdots, r$

$$(^*) \quad \omega_i(\nu, \theta) \geqq 0, \qquad \Delta_{10}\omega_i(\nu, \theta) \leqq 0, \qquad \Delta_{01}\omega_i(\nu, \theta) \leqq 0, \qquad \Delta_{11}\omega_i(\nu, \theta) \geqq 0$$

in R_1. But it is known that for such functions[29]

$$\lim \frac{1}{\pi^2} \int_0^w \int_0^z \omega_i(\nu, \theta) \frac{\sin \alpha\nu}{\nu} \frac{\sin \beta\theta}{\theta} \, d\nu \, d\theta = \tfrac{1}{4}\omega_i(+0, +0).$$

[29] Bochner ([8], p. 201) proves this result for functions which satisfy (*) in $\bar{R}$: $0 \leqq x < \infty$, $0 \leqq y < \infty$. If $\omega(x, y)$ satisfies (*) in R_1, set $\bar{\omega}(x, y) = \omega(x, y)$ in R_1 and $= 0$ elsewhere in $\bar{R}$; then $\bar{\omega}(x, y)$ satisfies (*) in $\bar{R}$.

Hence

$$\lim_{\alpha,\beta\to\infty} J_1(\alpha, \beta) = \frac{1}{4}\sum_{i=1}^{r} c_i\omega_i(+0, +0) = \tfrac{1}{4}\varphi(w - 0, z - 0).$$

Similarly by appropriate transformations, we can treat the other three integrals, since all rectangles are finite, and show that their limits are respectively

$$\tfrac{1}{4}\varphi(w + 0, z - 0), \qquad \tfrac{1}{4}\varphi(w - 0, z + 0), \qquad \tfrac{1}{4}\varphi(w + 0, z + 0).$$

Adding them together, we get $\bar{\varphi}_{0,0}(w, z)$. This concludes the proof of the theorem.

The following important corollary follows from Theorem 9.

COROLLARY. *If* $a_0 \geqq 0$, $b_0 \geqq 0$, *and* (a_0, b_0) *is inside the region of bounded convergence of* $\mathfrak{L}(\varphi; s, t)$, *the inversion formula is valid at every point* (w, z) *such that* (B) *and* (C) *are true.*[30]

4. In the above theorem and corollary, (B) and (C) may be replaced by stronger conditions. For example, each of the following conditions implies (B):

(B_1) $V_\varphi[w - \delta, w + \delta; 0, y] \leqq Me^{b_0 y}$ for all $y \geqq 0$, and some $\delta > 0$.

(Apply (6), and use the fact that $\varphi(x, y) \in V_0$.)

(B_2) $\varphi(x, y)$ is of bounded variation in x on $(w - \delta, w + \delta)$, uniformly in y for all $y \geqq 0$.
(Set $b_0 = 0$ in (B).)

(B_3) $\varphi(x, y)$ is a function of class H on the semi-infinite rectangle

$$[w - \delta \leqq x \leqq w + \delta; 0 \leqq y].$$

(Set $b_0 = 0$ in (B_1).) Similar conditions imply (C).

5. The following theorems give conditions that the inversion formula hold for all $w > 0, z > 0$. As in the introduction, $v(x, y)$, $v'(x, y)$, $v''(x, y)$ are functions equal to 0 if $x = 0$ or $y = 0$ and equal to $V_\varphi[0, 0; x, y]$, $V'_\varphi[0, x; y]$, $V''_\varphi[x; 0, y]$ respectively for $x > 0$, $y > 0$.

[30] The proof given here is based upon one variable proofs given by Hamburger ([15], p. 6) for the case $\rho = 0$ and by Widder ([25], p. 710) for the case $\rho > 0$. Durañona y Vedia and Trejo claim to have proved the inversion formula for $\rho_1 = \rho_2 = 0$ for any positive point in the region of bounded convergence of $\mathfrak{L}(\varphi; s, t)$—essentially, using only (A). There is an error in their proof, however, when they say ([13], p. 460) that (in the notation of this paper)

$$\int_M^\infty e^{-b\theta}\varphi(w - \nu, z + \theta)\frac{\sin\beta\theta}{\theta}\,d\theta$$

has bounded variation in ν on $(0, w)$ by Lemma 1. But application of this lemma would require that $(w - \nu, z + \theta)$ belong to class H on $0 \leqq \nu \leqq w$, $M \leqq \theta < \infty$ (essentially(B_3)), and in general nothing is known about the variation of $\varphi(x, y)$ in this infinite region. However, another proof of Theorem 19 for $\rho_1 = \rho_2 = 0$ using (B) and (C) as well as (A) could be constructed along the lines of their proof.

THEOREM 20. *If there exist real constants M, a_0, b_0 such that*

(D) $$v'(x, y) \leqq Me^{a_0x+b_0y} \text{ for all } x \geqq 0,\ y \geqq 0,$$

(E) $$v''(x, y) \leqq Me^{a_0x+b_0y} \text{ for all } x \geqq 0,\ y \geqq 0,$$

then for every $\rho_1 \geqq 0$, $\rho_2 \geqq 0$ the formula

(48) $$\bar{\varphi}_{\rho_1,\rho_2}(w, z) = \frac{-1}{4\pi^2}\int_{b-i\infty}^{b+i\infty}\int_{a-i\infty}^{a+i\infty} \frac{e^{sw+tz}f(s, t)}{s^{\rho_1+1}t^{\rho_2+1}}\,ds\,dt$$

$$[a > a_0,\ a \neq 0,\ b > b_0,\ b \neq 0]$$

is valid at every point (w, z) where $w > 0$, $z > 0$.

Proof. Since $|\varphi(x, y)| \leqq v'(x, y)$, (A) is true. For every $w > 0$, there exists a $\delta > 0$ such that $w - \delta > 0$. Then from (D),

$$V'_\varphi[w - \delta, w + \delta; y] \leqq V'_\varphi[0, w + \delta; y] = v'(w + \delta, y) \leqq Me^{a_0(w+\delta)+b_0y} = M_1e^{b_0y},$$

and this is (B). Similarly, from (E), (C) is true for all $z > 0$. Hence the theorem is true.

(As before, a_0 or b_0 or both might be 0 in D and E. For these special cases, it is possible to obtain simpler direct proofs.)

COROLLARY 1. *If there exist real constants M, a_0, b_0 such that*

(F) $$v(x, y) \leqq Me^{a_0x+b_0y}$$

for all $x \geqq 0$, $y \geqq 0$, then for every $\rho_1 \geqq 0$, $\rho_2 \geqq 0$ formula (48) *is valid for all $w > 0$, $z > 0$.*

Since $v'(x, y) \leqq v(x, y)$, $v''(x, y) \leqq v(x, y)$, the proof is immediate.

COROLLARY 2. *If $a_0 \geqq 0$, $b_0 \geqq 0$, and $\mathfrak{L}(\varphi; s, t)$ converges absolutely at (a_0, b_0),* (48) *is valid for all $w > 0$, $z > 0$.*[31]

6. In the proof of (II), Theorem 19, only condition (A) was used. In the proof of (I), condition (C) was used only to get a bound for

$$\int_{b-i\beta}^{b+i\beta} \frac{e^{tz}}{t^{\rho_2}}\left[\int_{z-\eta}^{z+\eta} e^{-ty}\varphi(x, y)\,dy\right]dt$$

where $x \geqq X_1 \geqq w + 1$, and condition (B) was used only to get a bound for the corresponding integral with respect to x and s. However, if $\rho_2 > 1$, this term is bounded as a result of (A) alone, for its absolute value is less than

$$\int_{z-\eta}^{z+\eta} |\varphi(x, y)|\, e^{b(z-y)}\left[\int_{-\beta}^{\beta}\frac{d\mu}{|b + i\mu|^{\rho_2}}dy\right] < M\cdot B\cdot\int_{z-\eta}^{z+\eta} e^{a_0x+b_0y+b(z-y)}\,dy = Re^{a_0x}.$$

[31] Biggeri ([6], pp. 23-31) establishes the corresponding inversion formula for Dirichlet series in the cases of absolute and regular convergence. He states Corollary 2 for his integral; the proof is analogous to that given for series. And since (D) and (E) give results somewhat analogous to regular convergence (see Theorems 4 and 5) Theorem 20 may be considered a generalization of his formula for regular convergence.

This is the same type of bound as that used in the proof of (I); hence the following theorem is true.

THEOREM 21. *If* (A) *is true and* $\rho_1 > 1$, $\rho_2 > 1$, *the inversion formula* (48) *is valid at each* $w > 0$, $z > 0$.

COROLLARY. *If* $a_0 \geqq 0$, $b_0 \geqq 0$ *and* (a_0, b_0) *is in the region of bounded convergence of* $\mathfrak{L}(\varphi; s, t)$, *for every* $\rho_1 > 1$, $\rho_2 > 1$, *the inversion formula is valid at each* $w > 0$, $z > 0$.

7. The following theorem gives a different type of condition:

THEOREM 22. *If sections of* $\mathfrak{L}(\varphi; s, t)$ *are uniformly bounded in* λ *and* μ, *where* $s = a_0 + i\lambda$, $t = b_0 + i\mu$, *and* $\rho_1 > 0$, $\rho_2 > 0$, *formula* (48) *holds for all* $w > 0$, $z > 0$ $(a_0 \geqq 0, b_0 \geqq 0)$.

Proof. Since $| L(X, Y; a_0, b_0) | < M$, by Theorem 9, (A) is valid, and hence the proof of (II) follows as before. From Theorem 18, given ϵ, there exist X_0, Y_0, such that

$$| L(X_2, Y_2; a + i\lambda, b + i\mu) - L(X_1, Y_1; a + i\lambda, b + i\mu) | < \epsilon$$

for $X_2 > X_1 \geqq X_0$, $Y_2 > Y_1 \geqq Y_0$ and all λ and μ, where $a > a_0$, $b > b_0$, $| a | \neq 0$, $| b | \neq 0$. Hence

$$| J(X_2, Y_2; \alpha, \beta) - J(X_1, Y_1; \alpha, \beta) | < \epsilon \left(\int_{-\alpha}^{\alpha} \frac{e^{aw}\, d\lambda}{| a + i\lambda |^{\rho_1+1}} \right) \left(\int_{-\beta}^{\beta} \frac{e^{bz}\, d\mu}{| b + i\mu |^{\rho_2+1}} \right)$$
$$< \epsilon \cdot A \cdot B \cdot e^{aw+bz},$$

where

$$A = \int_{-\infty}^{\infty} \frac{d\lambda}{| a + i\lambda |^{\rho_1+1}}, \qquad B = \int_{-\infty}^{\infty} \frac{d\mu}{| b + i\mu |^{\rho_2+1}}$$

since $\rho_1 > 0$, $\rho_2 > 0$. Hence (I) is established.

8. So far, we have considered inversion of $f(s, t)$ by means of a double infinite integral. Let us now consider inversion by means of repeated integrals which are integrals of the type

$$\int_{b-i\infty}^{b+i\infty} dt \int_{a-i\infty}^{a+i\infty} h(s, t)\, ds = \lim_{\beta \to \infty} \int_{b-i\beta}^{b+i\beta} dt \lim_{\alpha \to \infty} \int_{a-i\alpha}^{a+i\alpha} h(s, t)\, ds.$$

Then the proof reduces to the corresponding theorems for one variable ([25], pp. 708, 710).

THEOREM 23. *If* M, a_0, b_0 *are real constants such that*

(D) $\qquad v'(x, y) < Me^{a_0x+b_0y} \qquad$ *for all* $x \geqq 0$, $y \geqq 0$,

then for every $\rho_1 \geqq 0$, $\rho_2 \geqq 0$ *the formula*

$$(49)\qquad \bar{\varphi}_{\rho_1,\rho_2}(w, z) = \frac{-1}{4\pi^2}\int_{b-i\infty}^{b+i\infty} \frac{e^{tz}}{t^{\rho_2+1}}\,dt\cdot\int_{a-i\infty}^{a+i\infty} \frac{e^{sw}}{s^{\rho_1+1}}\cdot f(s, t)\,ds$$

$$[a > a_0,\, b > b_0,\, a \neq 0,\, b \neq 0]$$

is valid for each $w > 0$, $z > 0$.

Proof. Under the hypothesis, $\mathfrak{L}(\varphi;\, s,\, t)$ converges boundedly for $\sigma > a_0$, $\tau > b_0$, and

$$f(s, t) = \int_0^\infty e^{-sx}\,d_x k(x, t),$$

where

$$k(x, t) = \int_0^\infty e^{-ty}\,d_y \varphi(x, y)$$

and $k(x, t)$ has bounded variation in every $(0, x)$. Hence by the theorem for one variable,

$$\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \frac{e^{sw}}{s^{\rho_1}} f(s, t)\,ds = \bar{k}_{\rho_1}(w, t),$$

where

$$\bar{k}_0(w, t) = \tfrac{1}{2}[\varphi(w + 0, t) + \varphi(w - 0, t)],$$

$$\bar{k}_{\rho_1}(w, t) = \frac{1}{\Gamma(\rho_1)}\int_0^w (w - x)^{\rho_1-1} k(x, t)\,dx \quad \text{for} \quad \rho_1 > 0,$$

$$\frac{-1}{4\pi^2}\int_{b-i\infty}^{b+i\infty} \frac{e^{tz}}{t^{\rho_2+1}}\,dt\cdot\int_{a-i\infty}^{a+i\infty} \frac{e^{sw}}{s^{\rho_1+1}} f(s, t)\,ds = \frac{1}{2\pi i}\int_{b-i\infty}^{b+i\infty} \frac{e^{tz}}{t^{\rho_2+1}}\,\bar{k}_{\rho_1}(w, t)\,dt.$$

Since $\varphi(w + 0, y)$ is of bounded variation in every $(0, y)$ and its variation in $(0, y)$ is less than $Me^{a_0w+b_0y}$, from the case of one variable,

$$\int_0^\infty e^{-ty}\,d_y\varphi(w + 0, y)$$

converges absolutely for $\tau > b_0$ and it is equal to $k(w + 0, y)$. Using this and the corresponding result for $k(w - 0, y)$, we get

$$\bar{k}_0(w, t) = \int_0^\infty e^{-ty}\,d_y\,\tfrac{1}{2}[\varphi(w + 0, y) + \varphi(w - 0, y)],$$

where the integral converges absolutely. Hence by the inversion formulas for one variable,

$$\frac{1}{2\pi i}\int_{b-i\infty}^{b+i\infty} \frac{e^{tz}}{t^{\rho_2+1}}\,\bar{k}_0(w, t)\,dt = \bar{\varphi}_{0,\rho_2}(w, z).$$

Suppose $\rho_1 > 0$. Then

$$\bar{k}_{\rho_1}(w, t) = \frac{1}{\Gamma(\rho_1)} \int_0^w (w - x)^{\rho_1 - 1} dx \cdot \int_0^\infty e^{-ty} d_y \varphi(x, y)$$

$$= \frac{t}{\Gamma(\rho_1)} \int_0^w (w - x)^{\rho_1 - 1} dx \cdot \int_0^\infty e^{-ty} \varphi(x, y)\, dy.$$

In the last member, the inner integral converges uniformly with respect to x; hence

$$\bar{k}_{\rho_1}(w, t) = t \int_0^\infty e^{-ty} A(y)\, dy,$$

where

$$A(y) = \frac{1}{\Gamma(\rho_1)} \int_0^w (w - x)^{\rho_1 - 1} \varphi(x, y)\, dx.$$

It is easily seen that $A(y)$ is of bounded variation in every $(0, y)$ and

$$| A(y) | < \bar{M} e^{b_0 y}$$

for all $y \geqq 0$. Thus

$$\bar{k}_{\rho_1}(w, t) = \int_0^\infty e^{-ty} d_y A(y)$$

and hence

$$\frac{1}{2\pi i} \int_{b - i\infty}^{b + i\infty} \frac{e^{tz}}{t^{\rho_2 + 1}} \bar{k}_{\rho_1}(w, t)\, dt = \bar{\varphi}_{\rho_1, \rho_2}(w, z).$$

The following result is obtained by a similar proof.

THEOREM 24. *If M, a_0, b_0 are real constants such that*

(E) $\qquad v''(x, y) < M e^{a_0 x + b_0 y}$ *for all* $x \geqq 0$, $y \geqq 0$,

then for every $\rho_1 \geqq 0$, $\rho_2 \geqq 0$

$$\bar{\varphi}_{\rho_1, \rho_2}(w, z) = \frac{-1}{4\pi^2} \int_{a - i\infty}^{a + i\infty} \frac{e^{sw}}{s^{\rho_1 + 1}} ds \cdot \int_{b - i\infty}^{b + i\infty} \frac{e^{tz}}{t^{\rho_2 + 1}} f(s, t)\, dt \tag{50}$$

$$[a > a_0,\, b > b_0,\, a \neq 0,\, b \neq 0]$$

for each $w > 0$, $z > 0$.

Combining Theorems 20, 23, 24, we have the result:

THEOREM 25. *If* (D) *and* (E) *are true,* $\bar{\varphi}_{\rho_1, \rho_2}(w, z)$ *is given by any one of the formulas* (48), (49), (50), *for every* $\rho_1 \geqq 0$, $\rho_2 \geqq 0$.

COROLLARY. *If*

$$v(x, y) < Me^{a_0x+b_0y} \text{ for all } x \geqq 0,\ y \geqq 0, \tag{F}$$

$\tilde{\varphi}_{\rho_1,\rho_2}(w, z)$ *is given by any one of formulas* (48), (49), (50).[32]

8. In this paper, inversion of the double Laplace integral has been treated by the classical method. It is also possible to obtain inversion formulas in terms of differential operators as has been done by Widder and others in the one variable case. This problem and the problem of necessary and sufficient conditions that $f(s, t)$ must satisfy in order that it be expressed by a Laplace integral will be considered in a forthcoming paper.

BIBLIOGRAPHY

1. C. R. ADAMS, *Note on multiple Dirichlet and multiple factorial series*, Annals of Mathematics, (2), vol. 33(1932), pp. 406–412.
2. C. R. ADAMS, *Kojima on double series*, Bulletin of the American Mathematical Society, vol. 39(1933), pp. 979–981.
3. C. R. ADAMS AND J. CLARKSON, *On definitions of bounded variation for functions of two variables*, Transactions of the American Mathematical Society, vol. 35(1933), pp. 824–854.
4. C. R. ADAMS AND J. CLARKSON, *Properties of functions $f(x, y)$ of bounded variation*, ibid., vol. 36(1934), pp. 711–730.
5. C. BIGGERI, *Sobre las series e integrales dobles de Dirichlet*, Revista de la Academia de Ciencias Exactas, Físico-Químicas y Naturales de Madrid, vol. 29(1933), pp. 311–326.
6. C. BIGGERI, *Sobre las series e integrales dobles de Dirichlet*, (con't.), ibid., vol. 30(1933), pp. 9–50.
7. C. BIGGERI, *Sur les séries et les intégrales multiples de Dirichlet*, Reale Istituto Lombardo di Scienze e Letteri [Milan], Rendiconti, (2), vol. 66(1933), pp. 809–818.
8. S. BOCHNER, *Vorlesungen über Fouriersche Integrale*, Leipzig, 1932.
9. B. CAMP, *Multiple integrals over infinite fields, and the Fourier multiple integral*, American Journal of Mathematics, vol. 39(1917), pp. 311–334.
10. C. CARATHÉODORY, *Vorlesungen über Reelle Funktionen*, 2d ed., Leipzig, 1927.
11. J. CLARKSON, *On double Riemann-Stieltjes integrals*, Bulletin of the American Mathematical Society, vol. 39(1933), pp. 929–936.
12. A. DURAÑONA Y VEDIA AND C. A. TREJO, *Recintos de convergencia de las integrales dobles de Laplace-Stieltjes*, Publicaciones de la Facultad de Ciencias Fisicomatematicas, Universidad Nacional de La Plata, Contribucion al Estudio de las Ciencias Fisicas y Matematicas, Series Matematica, vol. 1(1936), pp. 315–327.
13. A. DURAÑONA Y VEDIA AND C. A. TREJO, *Sobre la inversion de la integral doble de Laplace-Stieltjes*, ibid., vol. 1(1937), pp. 451–464.
14. M. FRECHET, *Extension au cas des intégrales multiples d'une définition de l'intégrale due à Stieltjes*, Nouvelles Annales de Mathématiques, (4), vol. 10(1910), pp. 241–256.
15. H. HAMBURGER, *Über eine Riemannsche Formel aus der Theorie der Dirichletschen Reihen*, Mathematische Zeitschrift, vol. 6(1920), pp. 1–10.
16. T. H. HILDEBRANDT, *Definitions of Stieltjes integrals of the Riemann type*, American Mathematical Monthly, vol. 45(1938), pp. 265–277.

[32] Vignaux ([23], p. 77) gives formula (50) in the region of absolute convergence for $\rho_1 = \rho_2 = 0$.

17. T. H. HILDEBRANDT AND I. J. SCHOENBERG, *On linear functional operations and the moment problem for a finite interval in one or several dimensions*, Annals of Mathematics, (2), vol. 34(1933), pp. 317–328.
18. E. W. HOBSON, *Theory of Functions of a Real Variable*, 3d ed., vol. 1, Cambridge, 1927.
19. E. W. HOBSON, *Theory of Functions of a Real Variable*, 2d ed., vol. 2, Cambridge, 1926.
20. J. HYSLOP, *A theorem on the integral of Stieltjes*, Proceedings of the Edinburgh Mathematical Society, vol. 44(1926), pp. 79–84.
21. T. KOJIMA, *On the double Dirichlet series*, Science Reports of the Tohôku Imperial University, (1) (Math., Phys., Chem.), vol. 9(1920), pp. 351–400.
22. F. LEJA, *Sur les séries de Dirichlet doubles*, Comptes-rendus du I Congrès des Mathématiciens des Pays Slaves, Warsaw, 1929, pp. 140–158.
23. J. C. VIGNAUX, *Sobre la transformación de Abel-Laplace de dos variables*, Anales de la Sociedad Científica Argentina, vol. 116(1933), pp. 76–78.
24. J. C. VIGNAUX, *Sobre un teorema de las integrales dobles de Abel-Laplace*, ibid., vol. 116(1933), pp. 289–295.
25. D. V. WIDDER, *A generalization of Dirichlet's series and of Laplace's integrals by means of a Stieltjes integral*, Transactions of the American Mathematical Society, vol. 31(1929), pp. 694–743.
26. W. H. YOUNG, *Multiple Fourier series*, Proceedings of the London Mathematical Society, (2), vol. 11(1912), pp. 141–184.
27. W. H. YOUNG, *On multiple integration by parts and the second theorem of the mean*, ibid., vol. 16(1917), pp. 273–293.
28. W. H. YOUNG, *On multiple integration*, Proceedings of the Royal Society (A), vol. 93(1917), pp. 28–41.

UNIVERSITY OF WISCONSIN.

NON-COMMUTATIVE CHAINS AND THE POINCARÉ GROUP

By William W. Flexner

J. W. Alexander, W. Mayer, A. W. Tucker and S. Lefschetz[1] have abstracted from convex complexes an algebraic system carrying a "homology theory" which, when the algebraic system is itself a complex, becomes the ordinary homology theory of the complex. Chains are there commutative and the elements of a homology group are classes of cycles, each class being composed of cycles whose difference bounds a chain of higher dimension. It is the object of this paper to abstract from finite convex complexes in another direction to obtain an algebraic system S carrying non-commutative chains and a "homology group" π of dimension 1 whose elements are classes of cycles whose differences "bound" non-commutative chains of higher dimension. "Subdivision" of this algebraic system will be defined; the group π will be shown to be invariant under this subdivision; and, a non-trivial step in this case, it will be shown that when S is a convex complex, π is the Poincaré group.

1. The system S consists of "cells" each having associated with it an integer called its dimension and a function F (meaning boundary) whose domain is S and whose range is a subset of the "chains" of S. The cells comprise the neutral cell 1 and n-dimensional cells $\{E^n_i\}$ (called n-cells) in finite number for $n = 0, 1, 2$. It is convenient to suppose that 1 is an n-cell for each n. To simplify the notation, zero- and one-cells (or their inverses) will often be denoted by O, T, U, V and a, b, x respectively.

By an n-chain will be meant a "word" in the sense of the theory of non-Abelian groups with a finite number of generators, the letters of the word being n-cells or their inverses. For instance,

$$(1.1)\qquad \begin{aligned} C^n &\equiv (E^n_{i_1})^{x_1} \cdots (E^n_{i_s})^{x_s}, \qquad x_k = \pm 1, \qquad \text{and} \\ D^n &\equiv (E^n_{j_1})^{y_1} \cdots (E^n_{j_t})^{y_t}, \qquad y_k = \pm 1 \end{aligned}$$

are n-chains, and if

$$C^nD^n = (E^n_{i_1})^{x_1} \cdots (E^n_{i_s})^{x_s}(E^n_{j_1})^{y_1} \cdots (E^n_{j_t})^{y_t},$$

and $(E^n_i)(E^n_i)^{-1}$ are written 1, the n-chains form a free group $\mathfrak{C}^n$. Obviously, more than one word defines the same element of $\mathfrak{C}^n$. This distinction between the *word* C^n and the *element* C^n of $\mathfrak{C}^n$ must be kept clear. A *normal form* for an element of $\mathfrak{C}^n$ can be obtained from a word giving rise to that element by sup-

Received February 21, 1941.

[1] S. Lefschetz, Bull. Am. Math. Soc., vol. 43(1937), pp. 345–359. (References to the other authors will be found on p. 345.)

pressing all adjacent pairs $(E_i^n)(E_i^n)^{-1}$ and is independent of the order in which the suppression is made.[2] An element in normal form will be called *normal.* Two elements are the same if and only if they have the same normal form. If two words are identical as words, the notation will be $A \equiv B$; if they define the same element of $\mathfrak{C}^n$, $A = B$.

The function F may now be defined.

$$F(C^n) \equiv (FE_{i_1}^n)^{x_1} \cdots (FE_{i_s}^n)^{x_s}. \tag{1.2}$$

(The definition of FE_i^2 will be found below, after Lemma 1.1.)

$$F1 \equiv 1, \tag{1.3}$$

$$FE_i^0 \equiv 1 \text{ for all } i, \tag{1.4}$$

$$FE_i^1 \equiv E_{i_1}^0(E_{i_2}^0)^{-1}, \text{ where } i_1, i_2 \text{ depend on } i. \tag{1.5}$$

The *outset* of $(E_i^1)^x$ is $E_{i_1}^0$ if $x = 1$, $E_{i_2}^0$ if $x = -1$.
The *finish* of $(E_i^1)^x$ is $(E_{i_2}^0)^{-1}$ if $x = 1$, $(E_{i_1}^0)^{-1}$ if $x = -1$.

The outset (finish) of the word C^n is the outset (finish) of its first letter (last letter).

A word C^1 (see (1.1)) has *property* ϕ if it $\equiv 1$ or if the finish of $(E_{i_j}^1)^{x_j}$ is the inverse of the outset of $(E_{i_{j+1}}^1)^{x_{j+1}}$ for $j = 1, 2, \cdots, s - 1$.

A word C^1 has *property* θ if it $\equiv 1$ or if it has property ϕ and the finish of $(E_{i_s}^1)^{x_s}$ is the inverse of the outset of $(E_{i_1}^1)^{x_1}$.

If $C = a_1 \cdots a_p$, then $a_{q+1}a_{q+2} \cdots a_r$, $0 \leqq q < r \leqq p$, is a *factor* of the word C.

LEMMA 1.1. *If C has property ϕ (or θ) and D is the normal form of C, then D has property ϕ (or θ) and the outset of D is the outset of C.*

The proof is immediate.

Returning to the definition of F, we define FE_i^2 for each i to be a 1-chain having property θ. It follows that $FFE_i^2 = 1$.

A single additional condition will be put on the system S.

Condition Γ. For every pair E_i^0, E_j^0, $i \neq j$ there is a 1-chain having property ϕ, outset E_i^0 and finish $(E_j^0)^{-1}$.

To define π fix a zero-cell 0. Then a 1-chain will be called a *cycle* if some word defining it has property θ with outset 0. The cycles form a subgroup $\mathfrak{Z}$ of $\mathfrak{C}^1$. The elements C^1 of $\mathfrak{C}^1$ such that $FC^2 = C^1$ for some 2-chain C^2 make up a subgroup $\mathfrak{F}$ of $\mathfrak{C}^1$. Let $\mathfrak{G}$ be the smallest invariant subgroup of $\mathfrak{C}^1$ containing $\mathfrak{F}$. Then the intersection $\mathfrak{H} = \mathfrak{Z} \cap \mathfrak{G}$ is an invariant subgroup of $\mathfrak{Z}$. The group π is defined as the factor group

$$\pi = \mathfrak{Z}/\mathfrak{H},$$

and is called the Poincaré group of S.

[2] K. Reidemeister, *Einführung in die Kombinatorische Topologie*, Braunschweig, 1932, p. 33.

LEMMA 1.2. *If $C \in \mathfrak{G}$ then*

$$C = A_1X_1A_1^{-1} \cdots A_tX_tA_t^{-1}, \tag{1.6}$$

where $A_i \in \mathfrak{C}^1$, $X_i \equiv (FE^2_{k_i})^{y_i}$ for some k_i and $y_i = \pm 1$ and conversely.

Every element C given by (1.6) is contained in $\mathfrak{G}$, so the set of all such elements is a subgroup $\mathfrak{G}_0$ of $\mathfrak{G}$. Since, for example, $ABXB^{-1}CX'C^{-1}A^{-1} = (ABXB^{-1}A^{-1})(ACX'C^{-1}A^{-1})$, $\mathfrak{G}_0$ is an invariant subgroup. As it contains $\mathfrak{F}$, $\mathfrak{G}_0$ is equal to $\mathfrak{G}$.

LEMMA 1.3. *The group π is independent of the particular choice of the zero-cell* 0.[3]

2. If S is a system, then $\bar{S}$ is a system and an *elementary subdivision* of S if $\bar{S}$ is related to S in either of the two ways I and II.

I. The elements of $\bar{S}$ are obtained from those of S by replacing a single 1-cell, a, of S by two 1-cells, a' and a'', and a zero-cell T^{-1}, where if $S \ni E^1 \neq a$, FE^1 in $\bar{S}$ is the same as FE^1 in S; while if $Fa \equiv UV$ in S, then

(2.1) $Fa' \equiv UT$ in $\bar{S}$,

(2.2) $Fa'' \equiv T^{-1}V$ in $\bar{S}$,

(2.3) $F(E^0) \equiv 1$ in $\bar{S}$ for every zero-cell,

(2.4) FE^2 in $\bar{S}$ is obtained by replacing a wherever it occurs in the FE^2 of S by $a'a''$.

That $\bar{S}$ satisfies the conditions of §1 for a system is immediate.

II. If E is a 2-cell of S and if

$$FE \equiv C_1DC_2 \text{ in } S, \tag{2.5}$$

where C_1, D, C_2 are 1-chains, then $\bar{S}$ is obtained from S by replacing E by two 2-cells E', E'' and a one-cell x, where

(2.6) if $S \ni E^2_i \neq E$, then FE^2_i in $\bar{S}$ is the same as FE^2_i in S;

(2.7) $Fx = FD$, Fx in normal form;

(2.8) $FE' \equiv C_1xC_2$;

(2.9) $FE'' \equiv x^{-1}D$;

(2.10) if $\bar{S} \ni E^1 \neq x$, then FE^1 in $\bar{S}$ is the same as FE^1 in S;

(2.11) if $\bar{S} \ni E^0$, $FE^0 = 1$ in $\bar{S}$.

Condition Γ is here trivially satisfied. Obviously $Fx \equiv E^0_i(E^0_j)^{-1}$ or 1, where E^0_i is the outset of D, and $(E^0_j)^{-1}$ is the finish of D, so FE' and FE'' have property θ. Hence S is a system.

[3] The proof of this lemma follows the same course as a similar proof in L. Pontrjagin, *Topological Groups*, translated by Emma Lehmer, Princeton, 1939, p. 220.

A system S' obtained from S by a finite succession of steps I and II is called a *subdivision* of S. It is immediate that if S represents a simplicial complex K, the regular subdivision of K is representable as an S' (see §4).

3. Let the groups for $\bar{S}$ corresponding to the groups $\mathfrak{C}^n, \cdots, \pi$ for S be called $\bar{\mathfrak{C}}^n, \cdots, \bar{\pi}$.

THEOREM 3.1. *If $\bar{S}$ is an elementary subdivision of S, then π and $\bar{\pi}$ are isomorphic.*

COROLLARY 3.1. *If S' is a subdivision of S, their Poincaré groups are isomorphic.*

For the fixed zero-cell $\bar{0}$ of $\bar{\pi}$ take $\bar{0} = 0$.

Proof of Theorem 3.1 *for* I. Make the following change of basis for the 1-chains of $\bar{S}$:

$$a' \to e(a'')^{-1}, \quad \text{where} \quad e = a'(a'');$$

$$E_j^1 \to E_j^1, \quad E_j^1 \in \bar{S}, \quad E_j^1 \neq a'.$$

Then $Fe = Fa$. After the change of basis the 1-cell a'' occurs trivially in 1-chains $\bar{C} \in \bar{\mathfrak{Z}}$ (i.e., occurs in combinations $a''(a'')^{-1}$ or $(a'')^{-1}a''$) because $\bar{C}$ has property θ, and T occurs only in Fa''. Let $\tau\bar{C}$ be the chain obtained from $\bar{C}$, after suppression of a'' by replacing e by a. Then τ effects a homomorphism $\bar{\mathfrak{Z}} \to \mathfrak{Z}$ such that every element of $\mathfrak{Z}$ is the image of some $\bar{C} \in \bar{\mathfrak{Z}}$. Showing that $\tau^{-1}\mathfrak{H} = \bar{\mathfrak{H}}$ will now suffice to prove Theorem 3.1 in case I.[4]

Let $F(E_{k_i}^2)^{y_i}$ be X_i in S and $\bar{X}_i$ in $\bar{S}$, $y_i = \pm 1$. Then $\tau\bar{X}_i = X_i$. But by Lemma 1.2, $C \in \mathfrak{H}$ implies that

$$C = A_1X_1A_1^{-1} \cdots A_tX_tA_t^{-1} \tag{3.0}$$

for some $k_1, \cdots, k_t$ and 1-chains A_i. Replacing a by e on the right side gives a chain $\bar{C}$ of $\bar{\mathfrak{H}}$ with $\tau\bar{C} = C$, so the inverse images of the elements of $\mathfrak{H}$ are in $\bar{\mathfrak{H}}$.

On the other hand, if $\bar{C} = A_1\bar{X}_1A_1^{-1} \cdots A_t\bar{X}_tA_t^{-1}$, only the A_i may contain a'' non-trivially. Since $\bar{C}$ is in $\bar{\mathfrak{Z}}$, a'' does not occur non-trivially in $\bar{C}$, so setting $a'' = 1$ on both sides of the equation leaves $\bar{C}$ unchanged, preserves the equality and frees A_i of a''.[5] Therefore $\tau\bar{C} \in \mathfrak{H}$ and every element of $\bar{\mathfrak{H}}$ is an inverse of some element of $\mathfrak{H}$, which, combined with the result of the preceding paragraph, gives $\tau^{-1}\mathfrak{H} = \bar{\mathfrak{H}}$.

Proof of Theorem 3.1 *in case* II. If $\bar{C} \in \bar{\mathfrak{C}}^1$, write

$$\bar{C} = x^{y_0}H_1x^{y_1} \cdots H_tx^{y_t}, \tag{3.1}$$

[4] B. L. van der Waerden, *Moderne Algebra*, Berlin, 1930, vol. 1, p. 136.

[5] *If $C = D$ and the word C does not contain the letter b, and D' is obtained from D by setting $b = 1$, then $C = D'$.* Prove this by induction on the number, n, of times b occurs in D. It is true for $n = 0$, so assume it for $n < s$ and prove it for $n = s$. If $s > 0$, D contains b^{-1} at least once and there are chains $H = 1$, E and G such that $D = EbHb^{-1}G$ or $D = Eb^{-1}HbG$, so $D = EG = EHG$ where EHG has $n < s$, and hence $D = D'$.

where y_0 or $y_t = 0$ is allowed, and where H_i is a chain of $\bar{\mathfrak{C}}^1$ free of x and hence is also a chain of $\mathfrak{C}^1$. If D is as given by (2.5), define

$$\tau\bar{C} = D^{y_0}H_1D^{y_1}\cdots H_tD^{y_t}. \tag{3.2}$$

Let $\bar{C} = \bar{G}$ and suppose $\bar{G}$ normal. Then it was obtained from $\bar{C}$ by a succession of suppressions of pairs $(E_i^1)(E_i^1)^{-1}$. If $E_i^1 \neq x$, the same suppression can be made in $\tau\bar{C}$. If the pair is xx^{-1}, it gives rise to $DD^{-1} = 1$ in $\tau\bar{C}$. So $\tau\bar{C} = \tau\bar{G}$, and τ assigns a unique element of $\mathfrak{C}^1$ to each element of $\bar{\mathfrak{C}}^1$. If $\bar{C} \in \bar{\mathfrak{Z}}$, then $\tau\bar{C} \in \mathfrak{Z}$, since $Fx = FD$ by (2.7). Hence τ defines a homomorphism $\lambda\colon \bar{\mathfrak{Z}} \to \mathfrak{Z}$. Because $C \in \mathfrak{Z}$ implies $C \in \bar{\mathfrak{Z}}$, every element of $\mathfrak{Z}$ is $\tau\bar{C}$ for some $\bar{C} \in \bar{\mathfrak{Z}}$. Two lemmas, by proving that $\lambda^{-1}\mathfrak{H} = \bar{\mathfrak{H}}$, suffice to complete the proof of the theorem.

Lemma 3.1. *If $C \in \mathfrak{H}$ is given by the right side of* (3.2) *and $\bar{C}$ as given by* (3.1) *is in $\bar{\mathfrak{Z}}$, then $\bar{C} \in \bar{\mathfrak{H}}$.*

Lemma 3.2. *If $\bar{C}$ as given by* (3.1) *is in $\bar{\mathfrak{H}}$, then $\tau\bar{C} \in \mathfrak{H}$.*

Proof of Lemma 3.1. First remark that in any multiplicative group the expression

$$g = a_0b_1a_1\cdots b_sa_s \tag{3.3}$$

may be written

$$g = (h_1b_1h_1^{-1})\cdots(h_sb_sh_s^{-1})g', \tag{3.4}$$

where

$$h_1 = a_0\,,\ h_2 = a_0a_1\,,\ \cdots,\ h_s = a_0a_1\cdots a_{s-1}\,,\ g' = a_0a_1\cdots a_s\,. \tag{3.5}$$

Secondly, $C \in \mathfrak{H}$ implies $C \in \bar{\mathfrak{H}}$. To see this notice that C is given by (3.0). Now A_i is in $\bar{\mathfrak{C}}^1$ as well as in $\mathfrak{C}^1$, and unless $E_{k_i}^2 = E$, $\bar{X}_i = X_i$. If $E_{k_i}^2 = E$,

$$A_iX_iA_i^{-1} = A_i[(FE')C_2^{-1}(FE'')C_2]^{\nu_i}A_i^{-1} \in \bar{\mathfrak{H}}.$$

Thirdly, if $C \in \mathfrak{H}$ is given by (3.2), (2.8) gives $x = D(FE'')^{-1}$; so

$$\bar{C} = [D(FE'')^{-1}]^{y_0}H_1\cdots H_t[D(FE'')^{-1}]^{y_t}$$

which is in the form (3.3) with $g = \bar{C}$, $b_i = FE''$ or its inverse or 1, $a_i = H_i$ or D or D^{-1} or 1. Then $\bar{C}$ can be put in the form (3.4) with $g' = C \in \bar{\mathfrak{H}}$ (by the second remark). But $b_i \in \bar{\mathfrak{F}}$, $h_i \in \bar{\mathfrak{C}}^1$; so $h_ib_ih_i^{-1} \in \bar{\mathfrak{H}}$ and $\bar{C} \in \bar{\mathfrak{H}}$.

Proof of Lemma 3.2. Suppose $\bar{C}$ to be given by the right side of (3.0) with $\bar{X}_i$ replacing X_i, and suppose $\bar{C} \in \bar{\mathfrak{Z}}$. Then $\tau A_i \in \mathfrak{C}^1$ and unless $E_{k_i}^2 = E'$ or E'', $\tau\bar{X}_i = X_i$. But by (2.8) and (2.9) $\tau(FE') = FE$ and $\tau(FE'') = 1$, so $\tau\bar{C} \in \mathfrak{H}$.

4. If K is a connected convex complex with boundary operator ∂, let K^2 be the subcomplex of K composed of its 0-, 1- and 2-dimensional cells. Then if σ

is the classical Poincaré group, $\sigma(K) = \sigma(K^2)$. But K^2 gives rise to a system S on identifying cells with the same dimension in K^2 and S and translating

$$\begin{array}{lll} \partial(E_i^0) = 0 & \text{into} & F(E_i^0) \equiv 1, \\ \partial(E_i^1) = -E_{i_1}^0 + E_{i_2}^0 & \text{into} & F(E_i^1) \equiv (E_{i_1}^0)(E_{i_2}^0)^{-1}, \\ \partial(E_i^2) = \sum_{j=1}^{r} \eta_i^j E_j^1 & \text{into} & F(E_i^2) = (E_{j_1}^1)^{\eta_i^{j_1}} \cdots (E_{j_r}^1)^{\eta_i^{j_r}}, \end{array}$$

where $j_1, \cdots, j_r$ is a permutation of $1, \cdots, r$ which gives FE_i^2 property θ. Since E_i^2 is a convex polygon, such a permutation may always be found. The connectedness of K gives condition Γ of §2.

THEOREM 4.1. *If the convex complex K gives rise as just described to the system S, π is isomorphic to $\sigma(K)$.*

The point O generating σ may be taken to be the same as that generating π. Then the cycles of $\mathfrak{Z}$ identify themselves with the simplicial paths.[6] To prove Theorem 4.1 each element $C \in \mathfrak{H}$ must now be shown to correspond to a product of closed paths (cgc^{-1}), where c is a path from O to the vertex of a 2-cell, and g is a path from that vertex around the boundary of the 2-cell. This is a result of Theorem 4.2 below and from it Theorem 4.1 follows in the standard way.

Capital letters from now on denote 1-chains, small letters 1-cells.

LEMMA 4.1. *If in S*

$$V \equiv D_0U_1D_1 \cdots U_tD_t = Y$$

and

(i) *U_i has property ϕ and is normal,*

(ii) *Y is normal,*

(iii) *$V \equiv (a_1 \cdots a_r)(a_{r+1} \cdots) \equiv GH$ with $G = 1$ implies that both a_r and a_{r+1} are in some U_i,*

then the outset of V is the outset of Y.

Suppose $Y \equiv y_1 \cdots y_s$. Then $V \equiv G_0y_1G_1 \cdots y_sG_s$, where $G_i = 1$ for $i = 0, 1, \cdots, s$. Taking the G of the hypothesis to be G_0 and using (i) show that the finish of a_r is inverse to the outset α of y_1 which is also the outset of Y. Let $W_1 = 1$ be a factor of G_0 ending in a_r and such that the number of letters in the word W_1 is a minimum. Then $G_0 = W_0W_1$, $W_0 = W_1 = 1$ and $W_1 = a_k \cdots a_r$. If W_1 has no proper factors equal to 1, $a_k = (a_r)^{-1}$. If W_1 has proper factors equal to 1, these include neither a_k nor a_r (since W_1 is minimal). But W_1 can be reduced to 1 by successive suppression of proper factors that equal 1. Hence the next to the last step of the suppression leaves $W_1 = a_ka_r = 1$, and again $a_k = (a_r)^{-1}$. So α is the outset of a_k and the hypotheses of

[6] See H. Seifert and W. Threlfall, *Lehrbuch der Topologie*, Leipzig, 1934, p. 158, "Kantenwege", etc.

the lemma make it possible to go through the argument again with W_0 replacing G_0. In a finite number of steps this reasoning yields Lemma 4.1.

The converse of Lemma 1.1: "D has ϕ implies C has ϕ" is obviously false. Theorem 4.2 is a weaker substitute for this converse. As before let $F(E^2_{k_i})^{\pm 1} = X_i$, so each X_i is normal and has property θ. Furthermore let $\Lambda_i = A_i X_i A_i^{-1} \neq 1$, where A is an arbitrary normal word. Now consider elements Z of $\mathfrak{C}^1$ such that Z *has* property ϕ and

$$Z = W \equiv \Lambda_1 \Lambda_2 \cdots \Lambda_s \neq 1 \tag{4.1}$$

for some $\{X_i\}$ of the set of boundaries. Let

$$s(W) = s, \tag{4.2}$$

the number of factors Λ_i in W,

$$\gamma(W) = \sum_{i=1}^{s} \alpha(A_i), \tag{4.3}$$

where $\alpha(A_i)$ is the number of letters in the word which is the normal form of A_i, and let $\beta_k(W)$ be the number of factors Λ_i in W such that the first k letters of A_i form the same word as the first k letters of A_1. If $A_1 = 1$ or if $\alpha(A_1) < k$, then $\beta_k(W)$ shall be zero.

Among the W such that $W = Z$ let those with $s(W)$ a minimum form a class $\mathfrak{M}(Z)$. Let $\mathfrak{N}(Z)$ be the subset of $\mathfrak{M}(Z)$ consisting of those W for which $\gamma(W)$ is a minimum. Let $\mathfrak{S}^1$ be the subset of $\mathfrak{N}(Z)$ consisting of those W for which $\beta_1(W)$ is a maximum. Let $\mathfrak{S}^k$, $k > 1$, be the subset of $\mathfrak{S}^{k-1}$ consisting of those W for which $\beta_k(W)$ is a maximum. The set $\mathfrak{S}^{\gamma(W)}$ is non-empty for a Z satisfying (4.1). Call *standard* those $W \in \mathfrak{S}^{\gamma(W)}$.

THEOREM 4.2. *Given a* $Z \neq 1$ *with property* ϕ *and satisfying* (4.1) *there is a*

$$\overline{W} \equiv \overline{\Lambda}_1 \overline{\Lambda}_2 \cdots \overline{\Lambda}_s = Z$$

such that

(i) $\overline{\Lambda}_i = \bar{A}_i \bar{X}_i \bar{A}_i^{-1}$, $\bar{A}_i$ *normal*,
(ii) $\overline{\Lambda}_i$ *has property* θ *for each* i,
(iii) *the outset of* $\overline{\Lambda}_i$ *is the same for each* i *as the outset of the normal form of* Z,
(iv) $\overline{W} \in \mathfrak{N}(Z)$.

Putting $Z \in \mathfrak{Z}$, $c = A_i$, $g = X_i$ for each factor shows that Theorem 4.2 implies Theorem 4.1.

COROLLARY 4.2. *The normal form of* Z *has property* θ.

By Theorem 4.2, $W = Z$ has property θ, so the corollary follows from Lemma 1.1.

Theorem 4.2 depends on two additional lemmas.

LEMMA 4.2. *If*

(i) $W \equiv Z \neq 1$

is standard,

(ii) $A_1 \equiv B_1C_1$,

(iii) $W \equiv B_1GH$,

(iv) $(a_{t+1}a_{t+2} \cdots a_r)(a_{r+1} \cdots) \equiv GH$ *with* $G = 1$,

then both a_r *and* a_{r+1} *are in some* X_j.

If the break between a_r and a_{r+1} occurs in Λ_1 and Lemma 4.2 is false, let

$$A_1 \equiv BC, \qquad A_1^{-1} \equiv DE$$

and

$$G \equiv CX_1D = 1. \tag{4.4}$$

Eliminating C from $\Lambda_1 = BCX_1C^{-1}B^{-1}$ by means of (4.4) gives $\Lambda_1 = BD^{-1}X_1DB^{-1}$; so, since $W \in \mathfrak{N}(Z)$, $\alpha(BD^{-1}) \geqq \alpha(BC)$, even though BD^{-1} is not known to be normal. Therefore $\alpha(D) \geqq \alpha(C)$, $B^{-1} \equiv RE$ and $D \equiv C^{-1}R$. Then (4.4) becomes

$$G \equiv CX_1C^{-1}R = 1, \tag{4.5}$$

$\Lambda_1 = BE = E^{-1}R^{-1}E = E^{-1}CX_1C^{-1}E$ and $\alpha(E^{-1}C) \geqq \alpha(BC)$ so $\alpha(E) \geqq \alpha(B)$ and $R \equiv 1$. But then (4.5) gives $\Lambda_1 = 1$ and this contradicts $W \in \mathfrak{M}(Z)$. Hence if the break occurs in Λ_1, the lemma is true.

If the break occurs in Λ_j, $j > 1$, let

$$M \equiv \Lambda_2 \cdots \Lambda_{j-1},$$

$$N \equiv \Lambda_{j+1} \cdots \Lambda_s,$$

$$A_1 \equiv B_1C_1, \qquad A_j \equiv B_jC_j.$$

Then (4.1) becomes

$$W \equiv (B_1C_1X_1C_1^{-1}B_1^{-1})M(B_jC_jX_jC_j^{-1}B_j^{-1})N. \tag{4.6}$$

Now assume Lemma 4.2 false. This means that for some j either

$$G \equiv C_1X_1C_1^{-1}B_1^{-1}MB_jC_jX_jC_j^{-1} = 1 \tag{4.7}$$

or

$$G \equiv C_1X_1C_1^{-1}B_1^{-1}MB_j = 1. \tag{4.8}$$

Using (4.7) to eliminate B_1 from (4.6) gives

$$Z = M\Lambda_j(B_jC_1X_1C_1^{-1}B_j^{-1})N. \tag{4.9}$$

Using (4.7) to eliminate B_j in (4.6) gives

$$Z = B_1B_j^{-1}N, \tag{4.10}$$

$$Z = (B_1C_jX_jC_j^{-1}B_1^{-1})\Lambda_1MN \equiv V_1. \tag{4.11}$$

If (4.8) is used in the same ways,

$$(4.12) \qquad Z = M(B_jC_1X_1C_1^{-1}B_j^{-1})\Lambda_jN,$$

$$(4.13) \qquad Z = B_1C_jX_jC_j^{-1}B_j^{-1}N,$$

$$(4.14) \qquad Z = (B_1C_jX_jC_j^{-1}B_1^{-1})\Lambda_1MN \equiv V_2 .$$

Since (4.6) is standard, $W \in \mathfrak{N}(Z)$ and so either (4.9) or (4.12) gives $\alpha(B_jC_1) \geqq \alpha(B_1C_1)$; therefore $\alpha(B_1) \leqq \alpha(B_j)$. But either (4.11) or (4.14) gives $\alpha(B_1C_j) \geqq \alpha(B_jC_j)$, so $\alpha(B_1) \geqq \alpha(B_j)$. Hence neither (4.7) nor (4.8) can hold unless $\alpha(B_1) = \alpha(B_j)$ and B_1C_j is normal and V_1 and V_2 are in $\mathfrak{N}(Z)$.

It is impossible that $B_1 = B_j$, for then (4.10) yields $Z = N$ and (4.13) yields $Z = \Lambda_jN$. This contradicts the condition $W \in \mathfrak{M}(Z)$. But $B_1 \equiv 1$ and $\alpha(B_1) = \alpha(B_j)$ implies $B_j \equiv 1$ and $B_1 \equiv B_j$, so that

$$(4.15) \qquad B_1 \equiv 1 \text{ is impossible.}$$

But now in (4.11), (4.14) $\beta_{\alpha(B_1)}(V_a) > \beta_{\alpha(B_1)}(W)$ for $a = 1, 2$. This contradicts the standard character of W; and so the hypothesis that Lemma 4.2 is false is untenable.

LEMMA 4.3. *If* $W = Z$ *is given by* (4.1) *and is standard, then*

(a) *the outset of* Λ_1 *is the outset of* Z,

(b) Λ_1 *has property* θ.

If W did not satisfy the hypotheses of Lemma 4.1 with $V \equiv W$, $Y = Z$ and $U_i \equiv X_i$, then Lemma 4.2 would be false with $B_1 \equiv 1$ and by (4.15) this is impossible. Hence by Lemma 4.1 the outset of W (i.e., the outset of Λ_1) is the outset of Z.

If $A_1 \equiv 1$, (b) is obvious. If $A_1 \equiv a_1 \cdots a_p \neq 1$ and $X_1 \equiv a_{p+1}a_{p+2} \cdots$, suppose $B_1 = a_1 \cdots a_q$, $q \leqq p$, has property ϕ. Then by (a) $B_1^{-1}Z$ has property ϕ and Lemma 4.2 shows that the hypotheses of Lemma 4.1 hold with $Y = B_1^{-1}Z$ and $U_i \equiv X_i$. Hence the outset of $(a_q)^{-1}$ is the outset of a_{q+1}. Hence $a_1 \cdots a_qa_{q+1}$ has property ϕ if $a_1 \cdots a_q$ has. From this and the fact that X_i has property θ it follows that Λ_1 has property θ.

The proof of Theorem 4.2 now depends on an induction on the number of factors Λ_i in the right side of (4.1). If that number is one, the theorem is Lemma 4.3. Now assume the theorem when the number of factors is less than s and prove it for s factors. Let W be standard and set $Z' = \Lambda_1^{-1}Z = \Lambda_2 \cdots \Lambda_s \equiv W'$. Since $\Lambda_1^{-1} \equiv A_1X_1^{-1}A_1^{-1}$, the outset of Z' is the same as the outset of Z. Obviously $W' \in \mathfrak{N}(Z')$, so by hypothesis of induction there are $\bar{\Lambda}_2, \cdots, \bar{\Lambda}_s$ such that $\bar{\Lambda}_2 \cdots \bar{\Lambda}_s = Z'$ satisfies Theorem 4.2. Hence $Z = \Lambda_1\bar{\Lambda}_2 \cdots \bar{\Lambda}_s$ also satisfies that theorem.

CORNELL UNIVERSITY AND INSTITUTE FOR ADVANCED STUDY.

RINGS WITH MULTIPLE-VALUED OPERATIONS

By Robert S. Pate

Introduction

One of the most familiar generalizations of the concept of a "group" is that of a "multigroup" in which we retain essentially all of our postulates except that two elements combine into a subset of elements rather than a single element of the set. Since the concept of a group is fundamental in other divisions of algebra, it is natural to wonder just what one would obtain if the groups there were to be replaced by multigroups. This paper is an investigation into various parts of algebra in which this replacement has been made. It is essentially the material presented in a thesis written at the University of Illinois under the guidance of Professor Brahana. I am greatly indebted to him and to Professor Baer for many criticisms and suggestions on the material.

It shall be assumed that the reader has a general knowledge of the material on generalizations of groups.[1]

In §1 we introduce the concept of a generalized ring. This is defined much in the same way as is an ordinary ring, the fundamental difference between the two being that in our case addition and multiplication are not necessarily unique. We also obtain some elementary properties of those of our rings which are additive groups. An interesting result in this connection is that every product contains the same number n of elements, and n is a divisor of the order of the additive group.

We define an ideal which is similar to an ordinary ideal. By the use of a certain correspondence Q among the elements of our ring, we can isolate a certain subset of ideals composed of Q-ideals. We show that a satisfactory choice of the correspondence Q allows us to establish a representation of a Q-ideal which is similar to the classical one. It follows from our development, except in certain cases, that any structure considerations must be made in terms of these Q-ideals.

We consider the multiplicative properties of the elements in §4. We also discuss the case when we have unique factorization of elements into prime elements. This is done as preliminary work to §5.

In certain types of our rings we are able to define an "indeterminate domain" in which the product of two polynomials is a subset of polynomials. For some

Received March 1, 1941; in revised form, May 13, 1941.

[1] A good bibliography may be found in M. J. Dresher and O. Ore, *Theory of multigroups*, Amer. Jour. of Math., vol. 60(1938), pp. 705–733. Other papers include J. E. Eaton and Oystein Ore, *Remarks on multigroups*, Amer. Jour. of Math., vol. 62(1940), pp. 67–71; J. E. Eaton, *Associative multiplicative systems*, Amer. Jour. of Math., vol. 62(1940), pp. 222–232; Howard Campaigne, *Partition hypergroups*, Amer. Jour. of Math., vol. 62(1940), pp. 599–612.

specialized kinds of rings we then show that any polynomials contained in the product of two so-called "primitive polynomials" is primitive. We are also able to establish an analogue of the Gauss lemma. However, it results that the indeterminate domain of what we term an integral domain is an integral domain if and only if the integral domain is an ordinary integral domain.

1. General concepts.

DEFINITION 1. A set R of undefined elements between every two elements, say a and b, of which a product ab and a sum $a + b$ is defined shall be called a *ring* if

(a) R is a multigroup with respect to addition,

(b) R is an associative multiplicative system with respect to multiplication,

(c) $a(b + c) = ab + ac$ and $(b + c)a = ba + ca$ for any three elements a, b, and c of R.

DEFINITION 2. An element 0 of R which is an identity of the additive multigroup shall be called a *zero* of R. If 0 is also an additive scalar, it shall be called a *scalar zero* of R. A scalar zero whose product with each element of R is itself shall be called a *null zero*.

Let us consider the case when the additive multigroup of our ring is an ordinary group R. Then any element u of R defines an automorphism of R on the subgroup Ru. In particular, it defines an isomorphism between $R/H(u)$ and Ru/H, where H is an invariant subgroup of Ru and $H(u)$ is an invariant subgroup of R. Then $0u = H$ and ru is a coset of H under Ru. $u0$ and $0v$ are both invariant subgroups of R. Furthermore, they are both cosets of 00 in $R0$ and $0R$ respectively. Hence $u0 = 00 = 0v$ for any u and v. This shows us that H is independent of the choice of the element u of R, and is the same for both right and left multiplication.

A number of very useful results follow in a trivial way from the preceding discussion. We mention

THEOREM 1. *If a ring R is additively a finite group, every product contains the same number n of elements, and n is a divisor of the order of the additive group.*

In all subsequent discussion we shall assume that addition and multiplication are commutative.

DEFINITION 3. Any element 1 of a ring R shall be called a *unity element* if $1r \supset r$ for every element r of R.

DEFINITION 4. If R has at least one zero element 0, any element a' such that $a + a' \supset 0$ shall be termed an *additive inverse* of a. If R contains a unity element 1, any element b' such that $bb' \supset 1$ shall be called a *multiplicative inverse* of b.

Let n be any positive rational integer and a any element of R. The symbol na shall mean $a + a + a + \cdots n$ times, while a^n shall mean $aaa \cdots n$ times. We have $na + ma = (n + m)a$, $n(ma) = (nm)a$, $na + nb = n(a + b)$, $a^n a^m = a^{n+m}$, and $(a + b)^n$ equals the usual thing.

DEFINITION 5. A subring A of R is an *ideal* if $AR \subset A$.

Certain types of our ideals have interpretations in terms of the "isomorphisms" of our rings analogous to those that ordinary ideals have for ordinary rings. However, we shall not deal with them in this paper.

Definition 6. The *crosscut* $(A_1, A_2, \cdots)$ of any set $A_1, A_2, \cdots$ of ideals shall be the totality of elements of R contained in every A_i.

The crosscut of two submultigroups of a multigroup is not necessarily a submultigroup, because $a + x \supset b$ may have solutions in one submultigroup which are not in the other one. As a result, it can easily be shown that the crosscut of two ideals need not be an ideal. In order to avoid this situation we now introduce the concept of a Q-ideal. For our immediate use we shall need the notion of a T-set, which shall be any set of ideals of R such that the crosscut of any subset of them is an ideal of the set.

Definition 7. A correspondence among the elements of R shall be called a *Q-correspondence* if to a fixed element q and each element r of R correspond a subset $f(r)$ of elements x_i and another subset $g(r)$ of elements y_j such that $q + x_i \supset r$ and $q \subset y_j + r$ for all i and j. The x_i and y_j shall be called the *correspondents* of the element r.

Definition 8. An ideal A of R which contains q, and, if it contains r, contains $f(r)$ and $g(r)$ shall be called a *Q-ideal*.

Theorem 2. *The ideals of any T-set are Q-ideals for some Q, and for any Q the totality of Q-ideals form a T-set.*

Consider the crosscut of all the ideals of T. From this ideal we pick any element which we shall call q. Assume that no correspondence Q exists for q. Let r be some element of the ideal A of the T-set. Then for every expression $q + x \supset r$ there is an ideal A_x of T which does not contain x, but does contain r. The crosscut of all of the A_x is not an ideal, contrary to the assumption that we had a T-set.

We shall not give a proof of the second part of the theorem. The considerations that we have just discussed tell us that either some of the relationships $a + x \supset b$ have an infinite number of solutions, or we must proceed along the lines that we do.

Definition 9. The *union* of any set S of Q-ideals shall be the crosscut of all Q-ideals which contain every ideal of S. The union shall be indicated by brackets.

Theorem 3. *The Q-ideals of R form a structure.*

Definition 10. An ideal A shall be called *primary* if the only products bc, where b is not an element of A, which contain an element of A are from among those elements c such that there is a g such that c^g is contained in A. If g is always 1, A is a *prime* ideal.

2. Q-ideals. Hereafter, all ideals discussed shall be assumed to be Q-ideals.

Postulate I. *If $uv \subset A$, where A is an ideal, and if u' is any correspondent of u, then $u'v \subset A$.*

This is a postulate about Q.

Let A be any ideal and B any subset of elements of R. The symbol A/B shall denote the totality of elements e such that $Be \subset A$.

THEOREM 4. *The following two statements are equivalent*:

(a) *Q satisfies Postulate* I,

(b) *A/B is a Q-ideal for every A and B.*

For if ub and vb are in A, so are $(u + v)b$ and $(uv)b$. If Q satisfies Postulate I, all the correspondents of the various elements are in A/B, and $(ur)b = (ub)r \subset A$, so that A/B is a Q-ideal. The converse is clear when we consider those subsets B composed of but one element.

We shall not give the proof of

THEOREM 5. $(A, B)/C = (A/C, B/C)$.

For the remainder of this section we shall assume that Q satisfies Postulate I.

THEOREM 6. *Let A be any primary ideal. The totality B of elements c of R such that there is a g_c for which $c^{g_c} \subset A$ is a prime ideal.*

If $r \subset R$, we see that $(rc)^{g_c} = r^{g_c}c^{g_c} \subset A$. If b is any element of B, $(b + c)^{g_b+g_c} = \sum \binom{n}{d} b^{n-d}c^d \subset A$, where $n = g_b + g_c$. Postulate I tells us that if c' is any correspondent of c then $c'c^{g_c-1} \subset A$, and, since A is primary, there is a $g_{c'}$ such that $c'^{g_{c'}} \subset A$. Hence B is an ideal.

Now let $xy \supset z$, where $z \subset B$ and $x \not\subset B$. Hence, for some g_z, $z^{g_z} \subset A$. This implies that $x^{g_z-1}y^{g_z}$ contains an element of A. Continuing this process, we conclude that y is an element of B.

DEFINITION 11. Let A be any primary ideal and C an ideal containing A. If for every element c of C there is a g_c such that $c^{g_c} \subset A$, *C belongs* to A.

THEOREM 7. *A primary ideal A has only one prime ideal belonging to it.*

Let P and P' be two such primes. Then, if $p \subset P$, there is a g_p such that $p^{g_p} \subset A \subset P'$. Since P' is a prime, this means that $P \subset P'$. In the same way, $P' \subset P$.

DEFINITION 12. If $D = (D_1, D_2, \cdots)$, where the D_i are ideals, we shall call the set $D_1, D_2, \cdots$ a *representation* of D. If D is not the crosscut of fewer of the D_i, the representation shall be said to be *irreducible*. If the number of D_i is finite, we shall say that we have a *finite representation.*

Theorem 7 makes it easy to verify that we have

THEOREM 8. *If D is the crosscut of two primary ideals D_1 and D_2 which have the same prime P belonging to them, then D is primary and P belongs to D.*

THEOREM 9. *If, in any finite irreducible representation $D_1, D_2, \cdots, D_r$ of any ideal D by primary ideals, all the D_i have different primes P_i belonging to them, then D is not primary.*

Let P_i belong to D_i. Out of each P_i choose a_i, an element which is not in P_1, for all i except 1. For each a_i there is a g_i such that $a_i^{g_i} \subset D_i$. Hence, there is a g such that $D_1(\prod a_i)^g \subset D$. Therefore, if we choose d so that $d \subset D_1$, but $d \not\subset D$, while e is any element in $(\prod a_i)^g$, we know that $de \subset D$.

Now, assume that D is primary. Either $e \subset P$, or $d \subset D$. Since $d \not\subset D$, $e \subset P$, where P is the prime belonging to D. Since P is a prime, this implies that some a_i is in P. But $P \subset P_1$, and this a_i is in P_1, contrary to our assumption. Hence D is not primary.

THEOREM 10. *Let $D_1, \cdots, D_r$ be any primary ideals, and B any ideal. Then if B is not contained in any of the primes P_i belonging to the D_i, $(D_1, \cdots, D_r)/B = (D_1, \cdots, D_r)$.*

Let r be any element of R and b any element of B. Then if $rb \subset (D_1, \cdots, D_r)$, either $r \subset (D_1, \cdots, D_r)$ or b is contained in some P_i.

THEOREM 11. *If $D_1, \cdots, D_r$ and $D_1', \cdots, D_s'$ are two irreducible representations of the same ideal by primary ideals D_i and D_i' for which $P_i \neq P_j$ and $P_i' \neq P_j'$ when $i \neq j$, then $r = s$ and the primes belonging to the D_i are the same as those belonging to the D_i'.*

For some P_i, say P_1, $P_1 \not\subset P_i$ when $i \neq 1$. We have two possibilities: (1) P_1 is contained in some P_i', say P_1', or, (2) $P_1 \not\subset P_i'$ for any i. Let us first consider (1). We have two possibilities: (a) P_1 is the same as some P_i', or, (b) P_1 is contained in, but not equal to some P_i', say P_1'. If (a) is true, the first step of our proof is complete. If (b) is true, we can interchange P_1 and P_1' in what follows for (2). We have $(D_2, \cdots, D_r) = (R, D_2, \cdots, D_r) = (D_1/D_1, D_2/D_1, \cdots, D_r/D_1) = (D_1, D_2, \cdots, D_r)/D_1 = (D_1', D_2', \cdots, D_s')/D_1 = (D_1'/D_1, D_2'/D_1, \cdots, D_s'/D_1) = (D_1', D_2', \cdots, D_s') = (D_1, D_2, \cdots, D_r)$. Hence, we do not have an irreducible representation, contrary to assumption, and (1), (a) is true.

(D_1, D_1') is primary and P_1 belongs to it. If $r \subset R$ and $d \subset (D_1, D_1')$, $rd \subset D_1$ and $rd \subset D_1'$. Therefore, $D_1/(D_1, D_1') = R = D_1'/(D_1, D_1')$. Thus, $(D_2, \cdots, D_r) = (R, D_2, \cdots, D_r) = (D_1/(D_1, D_1'), D_2/(D_1, D_1'), \cdots, D_r/(D_1, D_1')) = (D_1, D_2, \cdots, D_r)/(D_1, D_1') = (D_1'/(D_1, D_1'), D_2'/(D_1, D_1'), \cdots, D_s'/(D_1, D_1')) = (D_2', \cdots, D_s')$. Obviously, we have a proof by induction.

These results are somewhat similar to some obtained for Dedekind structures. However, our results depend on multiplicative properties of the elements and are stronger than the results obtained there. Furthermore, we do not seem to have a Dedekind structure. This suggests that we should be able to obtain some decomposition results for structures which do not satisfy the Dedekind relation, probably for those satisfying some sort of an automorphism relationship. The classical development of ideal theory has been followed here.[2]

[2] The classical development may be found in van der Waerden, *Moderne Algebra* II.

3. The ascending chain condition. As in the preceding section, we shall restrict our attention to Q-ideals.

DEFINITION 13. $[a_1, a_2, \cdots]$ shall stand for the Q-ideal which is the crosscut of all the Q-ideals which contain every one of the elements of the *basis* $a_1, a_2, \cdots$. If the number of a_i is finite, we shall say that it is a *finite basis*.

In the customary way, we can prove that the necessary and sufficient condition that every ideal shall have a finite basis is that every ascending chain of Q-ideals shall have but a finite number of different members. This last condition, which is called the *ascending chain condition*, shall be assumed to obtain for the remainder of this section.

DEFINITION 14. An ideal A is *reducible* if there are ideals B and C such that $A = (B, C)$, where neither B nor C is equal to A. If A is not reducible, it is *irreducible*.

The ascending chain condition immediately gives us

THEOREM 12. *Every ideal may be written as the crosscut of a finite number of irreducible ideals.*

We now introduce

POSTULATE II. *If b is any element of R, Rb is an ideal.*

It is easy to verify

THEOREM 13. *Postulate* II *may be restated as follows: Let a, b, and c be any three elements of R such that $ab \supset c$. Then if c' is any correspondent of c, the relation $xb \supset c'$ has a solution for x in R.*

POSTULATE III. *If A and B are any two ideals, every element in $[A, B]$ occurs in an expression of the form $a + b$, where $a \subset A$ and $b \subset B$.*

If R contains a unity element, the subset Ra contains a, so that we have

THEOREM 14. *Let Q satisfy* II *and* III. *Then, if R contains a unity element, every element of $[a_1, a_2, \cdots, a_r]$ occurs in an expression of the form $\sum r_i a_i$, where r_i is an element of R.*

POSTULATE IV. *Let A be any ideal of R. Then if cd contains an element of A, cd is completely contained in A, for any two elements c and d of R.*

THEOREM 15. *Let R have a unity element and Q satisfy* I, II, III, *and* IV. *Then any additively reversible irreducible ideal is primary.*

Assume that the theorem is false for A. By Postulate IV, we know that there are elements c and d which are not in A such that $cd \subset A$, while there is no g such that $d^g \subset A$.

Suppose that $e_i \subset de_{i-1}$, where $e_0 = c$. Then, by Postulate I, the totality of elements r such that $re_i \subset A$ form an ideal, and we have an ascending chain because $A/e_i \subset A/e_{i+1}$. Hence, for some k, $A/e_k = A/e_{k+1}$. We shall show that A is reducible because $A = ([A, c], [A, Re_k])$, where e_k has been chosen so that it is not in A.

If $s \subset [A, c]$, Postulate III tells us that $s \subset a + rc$ and $sd \subset ad + r(cd)$. Hence, $sd \subset A$. If $s \subset [A, Re_k]$, $s \subset a' + r'e_k$ and $sd \subset a'' + r'(de_k)$. Since A is reversible, $r'(de_k)$ contains an element of A. Consequently, some $re_{k+1} \subset A$, and thus $re_k \subset A$. Therefore, $s \subset A$, and A is reducible, contrary to our assumption.

We now prove the main theorem of this section, namely,

THEOREM 16. *Let Q satisfy* I, II, III, *and* IV, *R have a unity element, and every irreducible ideal be additively reversible. Then every ideal may be expressed as the crosscut of finitely many primary ideals, no two of which have the same prime ideal belonging to them.*

By Theorem 12 every ideal may be written as the crosscut of finitely many irreducible ideals. Theorem 15 tells us that these irreducible ideals are primary. Theorem 8 indicates that we may replace each set of ideals in the representation which belong to the same prime ideal by their crosscut.

It is of interest to note that we now have the following partial Dedekind relationship.

THEOREM 17. *Let Q satisfy Postulate* III *and A be any additively reversible ideal. Then if $C \supset A$, $[A,(B, C)] = (C, [A, B])$.*[3]

Before leaving the subject of ideals and the correspondence Q, one should notice that some of the results apply to the subject of structures of submultigroups of a multigroup. Consequently, the correspondence Q is quite fundamental there also. It is apparent from a little thought that a complex and interesting theory of the relations between different Q correspondences and the various structures associated with them can be developed. However, it seemed more important to derive first an analogue of the decomposition of ordinary ideals. Since these considerations were quite lengthy, further development along this line at this time would probably become tedious.

4. **Integral domains.** We shall confine ourselves to commutative rings.

DEFINITION 15. A ring M with a scalar zero 0 and at least one unity element 1 shall be called an *integral domain* if

(1) $m_1m_2 \supset 0$ implies that either $m_1 = 0$ or $m_2 = 0$;

(2) $m_1m_2 \supset m_1$ implies that m_2 is a unity element, or $m_1 = 0$;

(3) $m_1m_2m_3 = mm_3$, where m is any element contained in m_1m_2.

DEFINITION 16. If $ab \supset c$, we shall say that a and b *divide* c. If c also divides a, then a and c are *associates*. Any element is *regular* if it has an associate which is a unity element.

It is clear that division is reflexive and transitive. If a and c are associates and $a \supset bc$, then b is a regular element or $c = 0$. Successive applications of (1), (2), and (3) of Definition 15 tell us that $01 = 0$, then $00 = 0$, and finally that $0a = 0$ for any a.

[3] The proof is essentially that used by Dresher and Ore in establishing a similar theorem about multigroups.

Let n and n' be any two elements of M and let $nn' \supset a$. Then $nn' = n(n'1) = (nn')1 = a1$. Thus we have proved

THEOREM 18. *The elements occurring in any product nn' are exactly those in $a1$, where a is any element in nn' and 1 is any unity element.*

DEFINITION 17. Let a and b be any two elements of M. Then any element c which divides both a and b and is divided by all elements which divide both a and b shall be called a gcd (a, b). If c is divided by both a and b, and if c divides any element which is divisible by both a and b, then c shall be called a lcm (a, b).

DEFINITION 18. An element which is not regular and whose only divisors are its associates and regular elements is a *prime*. The elements $a_1, a_2, \cdots$ are *relatively prime* if the only elements which divide all of them are regular elements.

In the customary way, we may show that we have

THEOREM 19. *If d is a gcd (a, b), $a \subset dt$, and $b \subset dt'$, then t and t' are relatively prime.*

Let a' be any additive inverse of a. Assume that m is any element of M, and that k and r are any two elements contained in ma and ma' respectively. Then, if $t \neq 0$, we have $t(k + r) = t(ma + ma') = tm(a + a') \supset tm0$. Hence, *if a' is any additive inverse of a, any element in ma' is an additive inverse of any element in ma.*

DEFINITION 19. Let F be an integral domain containing the integral domain M. Then F is a *quotient field* of M if

(1) 1 is a unity element of F and 0 is a zero of F,

(2) for every $m \neq 0$ of M there is an m^{-1} such that mm^{-1} contains a unity element,

(3) every element of F is contained in some product nm^{-1}, where m and n are elements of M.

M may be divided into disjunct subsets each of which consists of the elements in some product $1m$. In some cases, we may adapt the ordinary method of imbedding an ordinary integral domain in a field to these subsets.

If any integral domain M is contained in a quotient field F, and if m, n, and p are any three elements of M such that $mn^{-1} \supset p$ for any n^{-1}, then, since $m \subset mnn^{-1} = pn$, n and p are divisors of m.

If any integral domain M is contained in another integral domain with the same unity elements, any multiplicative inverse s^{-1} of a regular element s of M is itself contained in M. For, let s' be any multiplicative inverse of s which is contained in M. Then $s^{-1} \subset s^{-1}1 = s^{-1}ss' = 1s' \subset M$.

Let a, b, and c be any three elements such that $a + b \supset c$. If -1 is any additive inverse of 1, then $(-1)c \subset (-1)a + (-1)b$. Consequently, a', b', and c' exist which are additive inverses of a, b, and c respectively, such that $a' + b' \supset c'$.

DEFINITION 20. An integral domain is a *unique factorization domain* if

(1) $a \neq 0$ implies that there are primes p_i such that $a \subset \prod^n p_i$, or a is a regular element;

(2) if sets of primes p_i and q_i are such that $\prod^r p_i \supset a \subset \prod^s q_i$, then $r = s$ and there is a 1-1 correspondence p_i corresponding to $q_{j(i)}$ such that p_i and $q_{j(i)}$ are associates.

DEFINITION 21. Two such factorizations as referred to in (2) of the preceding definition shall be called *associated prime factorizations*. If the p_i and q_i are not primes, the factorizations shall be termed *associated*, provided that the rest of the conditions of (2) are satisfied.

The following facts may be shown in the customary way.

(a) The necessary and sufficient condition that two elements of a unique factorization domain be associated is that they have associated factorizations into primes.

(b) If the p_i are primes in a unique factorization domain, every element in $\prod^r p_i$ divides every element in $\prod^{r+s} p_i$.

(c) Let a and b be any two elements in a unique factorization domain such that a divides b. Then if $a \subset \prod^r p_i$, and $b \subset \prod^s q_i$, where the p_i and q_i are primes, then $r \leqq s$ and to every p_i is associated a $q_{j(i)}$.

(d) Any two non-zero elements in a unique factorization domain have at least one gcd in it.

(e) If a, b, and c are three elements in a unique factorization domain such that a and b are relatively prime to c, then ab is relatively prime to c.

(f) Any two quantities in a unique factorization domain M have a lcm in M.

5. Indeterminate domains. In this section we shall consider a commutative ring R with a null zero 0 and at least one unity element 1.

Consider the symbol x. With it and R associate the following conventions:

(1) $0x^n = x^n 0 = 0$.

(2) Form all possible new quantities $f(x) = a_0 + \sum^n a_i x^i$, where the a_i are subsets of elements of R and n is a positive integer. Then if $f_1(x) = b_0 + \sum b_i x^i$ and $f_2(x) = c_0 + \sum c_i x^i$ are any two of these quantities, we have the following rules:

(a) $f_1(x) = f_2(x)$ if $b_i = c_i$ for all i;

(b) $f_1(x) \supset f_2(x)$ if $b_i \supset c_i$ for all i;

(c) $f_1(x) + f_2(x) = (b_0 + c_0) + \sum (b_i + c_i)x^i$;

(d) $f_1(x)f_2(x) = \sum_i (\sum_k a_{i-k} b_k)x^i$.

If any $b_i x^i$ does not occur in $f_1(x)$, or $c_i x^i$ in $f_2(x)$, we agree to replace b_i, or c_i, by 0 in (c) and (d).

DEFINITION 22. The subset a_i in $f(x)$ of (2) shall be called the *coefficient* of x^i. We shall say that $f(x)$ is a *polynomial*, and if every a_i is a single element of R, a *polynomial with coefficients in R*. The totality of all polynomials with coefficients in R shall be denoted by $\bar{R}$.

It is clear that $\bar{R}$ is a ring containing R. Furthermore, 0 is a null zero of $\bar{R}$, the unity elements of R are exactly those of $\bar{R}$ and the regular elements of $\bar{R}$ are those of R.

THEOREM 20. *If $\bar{R}$ is an integral domain, R is an ordinary integral domain.*

In view of Theorem 1 and the fact that 0 is a null zero, if the theorem is false, two elements, say a and b, must exist such that $a + b$ contains at least two elements, say c and d. If $\bar{R}$ is an integral domain, $f_1f_2f_3 = f'f_3 = f''f_3$, where f' and f'' are any two elements of $\bar{R}$ which are in f_1f_2. Hence,

$$(1x + a)(1x + b)(1x + e) = (1x^2 + cx + r)(1x + e) = (1x^2 + dx + r)(1x + e),$$

where 1 is any unity element of R and r is any element in ab. But,

$$(1x^2 + cx + r)(1x + e) = 11x^3 + (1c + 1e)x^2 + (1r + ec)x + er,$$

and

$$(1x^2 + dx + r)(1x + e) = 11x^3 + (1d + 1e)x^2 + (1r + ed)x + er.$$

Consequently, $1c + 1e = 1d + 1e$, and $m1(c + e) = m(c + e) = m(d + e)$ for any m and e of R. Letting $e = 0$, we see that c and d are associates. Since R has no associates of 0, if w is any element of R, and w' is any additive inverse of w, $w + w' = 0$, and if u is any element of R, we can pick a v in R such that $w' + v = u$, and $w + w' + v + u = w$, i.e., $v + u = w$. Hence, w is an additive scalar. Since w is any element of R, R is an additive group. Hence, by Theorem 1, all the elements of R are multiplicative scalars.

We shall have occasion to refer to the following conditions on R:

(1) every non-regular element is divisible by at least one prime;
(2) if a and c are divisible by d, and $a + b \supset c$, b is divisible by d;
(3) if a prime p divides the elements of ab, it divides a or b;
(4) R is an integral domain;
(5) R is contained in a quotient field I which is an integral domain;
(6) R is a unique factorization domain.

DEFINITION 23. If R is an integral domain and f is any element of $\bar{R}$, then f is *primitive* if the only elements that divide all of its coefficients are the regular elements of R.

THEOREM 21. *Let R satisfy* (1), (2), (3), *and* (4). *Then if f_1 and f_2 are primitive, and if $f \subset f_1f_2$, f is primitive.*

Suppose that every coefficient of f is divisible by a non-regular element b. Then, every coefficient is divisible by some prime p. We may write $f_1 \subset f_1' + pf_1''$ and $f_2 \subset f_2' + pf_2''$, where f_1' is the sum of all the terms of f_1 whose coefficients are not divisible by p, and f_2' is formed similarly. Then, $f \subset f_1f_2 \subset f_1'f_2' + p(f_1'f_2'' + f_1''f_2' + pf_1''f_2'')$. Any element c in the coefficient of the highest power of x in $f_1'f_2'$ is not divisible by p by (3). Let any element in the coefficient of the same power of x in $p(f_1'f_2'' + f_1''f_2' + pf_1''f_2'')$ be d, and the corresponding one in f be r, so that $r \subset c + d$ for proper choice of c and d. By (2), c is divisible by p, contrary to fact. Hence the theorem is true.

Applying this last theorem to an R satisfying (1), (2), (3), and (4), we obtain two corollaries:

COROLLARY. *If a prime p divides every coefficient of f, and if $f \subset f_1f_2$, then p divides every coefficient of f_1 or f_2.*

COROLLARY. *Suppose that $f \subset f_1f_2$, and that c divides every coefficient of f. Then, if c has no non-regular divisors in common with every coefficient of f_1, c divides all the coefficients of f_2.*

THEOREM 22. *Let R satisfy* (2) *and* (4). *Then any element b is an additive inverse of any element a if and only if $b \subset 1'a$, where $1'$ is some additive inverse of any unity element.*

The remark following Theorem 19 tells us that every element in $1'a$ is an additive inverse of a. Now let b be any element such that $a + b \supset 0$. (2) tells us that a and b are associates, and consequently, $b \subset ar$ and $0 \subset a + b \subset a(1 + r)$. Hence, $1 + r \supset 0$.

DEFINITION 24. Let f be any element contained in $\bar{R}$. f is *reducible* in R if there are two elements f_1 and f_2 of $\bar{R}$ such that $f \subset f_1f_2$. f_1 and f_2 shall be called *factors* of f.

THEOREM 23. *Let R satisfy* (2), (5), *and* (6). *Let f be a primitive polynomial with coefficients in R, and such that there are polynomials g and h of $\bar{I}$ for which $f \subset gh$. Then, there are elements a and b of I and primitive polynomials g_0 and h_0 of $\bar{R}$ such that $g \subset ag_0$ and $h \subset bh_0$, and for any set a, b, g_0 and h_0 satisfying these conditions, ab contains only regular elements of R.*

Since I is a quotient field of R, we may write $g \subset \sum a_in_i^{-1}x^i$ and $h \subset \sum b_im_i^{-1}x^i$, where a_i, b_i, n_i, and m_i are elements of R, while n_i^{-1} and m_i^{-1} are multiplicative inverses of n_i and m_i. Let d_2 be a lcm of all the n_i. Let d_1 be a gcd of all the $a_in_i^{-1}d_2$. Then $g \subset d_2^{-1}d_1 \sum a_i'x^i$, where the a_i' are elements of R and are relatively prime as a totality. We may assume that d_1 and d_2 are relatively prime. Otherwise we should be able to replace them by appropriate elements in the product $d_2^{-1}d_1$. Similarly, $h \subset d_1'd_2'^{-1} \sum b_i'x^i$. Let $g_0 \subset \sum a_i'x^i$ and $h_0 \subset \sum b_i'x^i$, while f_0 we take to be any polynomial with coefficients in R such that $f \subset abf_0 \subset abg_0h_0$, where a is any element in $d_1d_2^{-1}$ and b is any element in $d_1'd_2'^{-1}$. Now $ab = (d_1d_1')d_2^{-1}d_2'^{-1} = rs^{-1}$, and $f \subset rs^{-1}f_0$. Hence, s divides r or has a gcd in common with the totality of coefficients of f_0. Since f_0 is primitive, s divides r. But r divides all the coefficients of f, and, consequently, is regular. Hence, s is regular and any element in ab is in R and is a regular element of R.

We shall say that a polynomial with coefficients in R is *monic* if the coefficient of the highest power of x is a unity element.

THEOREM 24. *Let R satisfy the conditions of Theorem 23. Then, if a monic polynomial of $\bar{R}$ is such that $f \subset gh$, where g and h are monic polynomials in $\bar{I}$, g and h belong to $\bar{R}$.*

By the preceding theorem, $g \subset ag_0$ and $h \subset bh_0$, where ab contains only regular elements of R and g_0 and h_0 are primitive polynomials of $\bar{R}$. Then, we

may write $a \subset cd^{-1}$, $b \subset qr^{-1}$, and we may assume that c is relatively prime to d and that q is relatively prime to r. Now dr divides cq, d divides q, and r divides c, so we write $q \subset dq_0$, $c \subset rc_0$, and $ab = cd^{-1}qr^{-1} = rc_0d^{-1}dq_0r^{-1} = c_0q_0$. Since ab contains only regular elements, c_0 and q_0 are regular. Hence we may choose g_0', h_0', a', and b' such that $a' \subset rd^{-1}$ and $b' \subset cr^{-1}$. The last coefficient of g_0 is contained in $rd^{-1}e$, where e is the last coefficient of g_0' and the last coefficient of h_0 is in $dr^{-1}e'$ analogously. Then, since g is monic, $rd^{-1}e$ contains a unity element, as does $dr^{-1}e'$. Hence, $re \supset d$ and $de' \supset r$. Thus r and d are associates and divide each other. Therefore, a' and b' belong to R, and $g \subset a'g_0'$, while $h \subset b'h_0'$, so that h and g belong to $\bar{R}$.

UNIVERSITY OF ILLINOIS.

INTRANSITIVE ABELIAN ALMOST-TRANSLATION GROUPS OF ALMOST-PERIODIC FUNCTIONS

By Fred Assadourian

It is our purpose here to generalize as completely as possible the results which H. Bohr and D. A. Flanders obtained in a recent paper[1] concerning the almost-translation groups of almost-periodic functions.

We deal with a finite set of m distinct one-valued, continuous almost-periodic functions $f_1(t), \cdots, f_m(t)$ (written collectively as $[f(t)]$) of a real variable, defined for $-\infty < t < +\infty$. The substitution S is defined over the integers $1, \cdots, m$, and if S takes j into k, then $S\{f_j(t)\} = f_k(t)$. For given $\epsilon > 0$ the real number τ is said to ϵ-perform the substitution S on $[f(t)]$ if

$$[\,| f_h(t + \tau) - Sf_h(t) \,| < \epsilon] \qquad (-\infty < t < +\infty;\, h = 1, \cdots, m).$$

A substitution S on $[f(t)]$ is defined as an *almost-translation substitution* if the set $\{\tau_{(S)}(\epsilon)\}$ of all real numbers τ each of which ϵ-performs the substitution S on $[f(t)]$ is relatively dense for every positive ϵ.

The set of almost-translation substitutions of $[f(t)]$ forms an Abelian group called the *almost-translation group* G of $[f(t)]$. Bohr and Flanders set up a list of necessary and sufficient conditions which a set $[f(t)]$ must satisfy in order that it have a given transitive Abelian group G as its almost-translation group, and then they went on to prove that any transitive group G can serve as the almost-translation group of some set $[f(t)]$. The authors left open the question of intransitive Abelian groups. We settle this question conclusively and arrive at results which cannot be generalized further since almost-translation groups must be Abelian substitution groups. We denote the main theorem of Bohr and Flanders by Theorem A and present it with changes in notation.

Theorem A.[2] *Let $[f(t)]$ be a finite set of m almost-periodic functions, ${}_1(t), \cdots, f_m(t)$; and let G be an arbitrary transitive Abelian group of m substitutions which we denote by*

$$S(1) = \begin{pmatrix} 1, 2, \cdots, m \\ 1, 2, \cdots, m \end{pmatrix}, \cdots, S(m) = \begin{pmatrix} 1, 2, \cdots, m \\ m, \cdots \end{pmatrix}.$$

Then in order that $[f(t)]$ be composed of distinct functions and have G as its almost-translation group, it is necessary and sufficient that the following four conditions be fulfilled:

Received March 10, 1941.

[1] *Algebraic equations with almost-periodic coefficients*, Kgl. Danske Videnskabernes Selskab, Mathematisk-fysiske Meddelelser, vol. 15(1937), pp. 1–49.

[2] Loc. cit., p. 32.

(i) *All the functions $f_h(t)$ have exactly the same Fourier exponents λ_n, and the absolute values of the corresponding Fourier coefficients are the same.*

(ii) *Further, the Fourier series of the functions $f_h(t)$ have the form*

$$f_h(t) \sim \sum_n \chi_n[S(h)]a_n \exp (i\lambda_n t),$$

where $\chi_n[S]$, for each n, is a character of the group G.

(iii) *Those characters $\chi_n[S]$ which actually occur form a generating system of the character-group G^* of G.*

(iv) *If any finite set $\lambda_1, \cdots, \lambda_N$ of Fourier exponents is connected by a linear relation*

$$g_1\lambda_1 + \cdots + g_N\lambda_N = 0$$

with integral coefficients, then the corresponding characters are connected by the relation

$$\{\chi_1[S(h)]\}^{g_1} \cdots \{\chi_N[S(h)]\}^{g_N} = 1 \qquad (1 \leqq h \leqq m).$$

Before generalizing this theorem we introduce a system of notation. Let G now be an intransitive Abelian group of substitutions on the integers $1, \cdots, m$ which split up into the following *systems of intransitivity*: $1, \cdots, m_1$; $m_1 + 1, \cdots, m_1 + m_2$; $\cdots$. We denote the number of integers in the k-th system by m_k and the number of systems by r. The substitutions of G operating just on the k-th system form a transitive Abelian group $G(k)$. Therefore, Theorem A can be applied to these systems of intransitivity.

Now we suppose that the finite set of m almost-periodic functions $[f(t)]$ has G as its almost-translation group. We group these functions into their systems of intransitivity in the following way:

$$f_{k1}, f_{k2}, \cdots, f_{km_k},$$

where $k = 1, \cdots, r$ and $\sum m_k = m$.

We write the effect of any substitution S in G on the subset $[f_{kj}]$, k fixed, as $S(k, h)$, where the latter denotes one of the m_k transitive substitutions of $G(k)$. Also, we may assume[3] that the functions in the k-th subset have been arranged so that $S(k, h)(f_{k1}) = f_{kh}$. We denote the character-group of $G(k)$ by $G^*(k)$. Its characters will be denoted by $\chi_{kj}[S(k, h)]$ $(h, j = 1, \cdots, m_k)$. We write any S in G as $S = \prod_k S(k, h)$, where h depends on k.

For an intransitive Abelian group G, Theorem A generalizes into

THEOREM 1. *Let $[f(t)]$ be a finite set of m almost-periodic functions $f_1(t), \cdots, f_m(t)$; and let G be an arbitrary intransitive Abelian group on the integers $1, \cdots, m$ decomposable into r systems of intransitivity, each consisting of m_k integers, $k = 1, \cdots, r$, $\sum m_k = m$. Then, in order that $[f(t)]$ be composed of distinct functions and have G as its almost-translation group, it is necessary and sufficient that the following conditions be fulfilled:*

[3] Loc. cit., p. 27.

(i) *For fixed k, all functions* $f_{kh}(t)$ $(h = 1, \cdots, m_k)$ *have the same Fourier exponents* λ_{kn} *with non-vanishing coefficients, and the absolute values of the corresponding Fourier coefficients are the same.*

(ii) *The Fourier series of the function* $f_{kh}(t)$ *has the form*

$$f_{kh}(t) \sim \sum_n \chi_{kn}[S(k, h)]\, a_{kn} \exp(i\lambda_{kn} t), \tag{1}$$

where, for each n, $\chi_{kn}[S(k, h)]$ *is a character of* $G(k)$.

(iii) *Those characters for fixed k which actually occur in* (1) *form a generating system of* $G^*(k)$.

(iv) *For two unequal values of k no one of the series* (1) *corresponding to one value of k is identical with any of the series* (1) *corresponding to the other value of k.*

(v) *If any finite set* $\lambda_{11}, \cdots, \lambda_{1N_1}; \cdots; \lambda_{r1}, \cdots, \lambda_{rN_r}$ *of the Fourier exponents of* $[f(t)]$ *is connected by a linear relation*

$$\sum_{k,n} g_{kn}\lambda_{kn} = 0 \tag{2}$$

with integral coefficients, then the corresponding characters are connected by the relation

$$\prod_{k,n} \{\chi_{kn}[S(k, h)]\}^{g_{kn}} = 1 \tag{3}$$

for every S in G. If $S = \prod_k S(k, h)$ *is a substitution on the systems of intransitivity of G but is not in G, then there must exist a set of integers* $g_{11}, \cdots, g_{1M_1}; \cdots; g_{r1}, \cdots, g_{rM_r}$ *such that*

$$\sum_{k,n} g_{kn}\lambda_{kn} = 0 \tag{4}$$

and

$$\prod_{k,n} \{\chi_{kn}[S(k,h)]\}^{g_{kn}} \neq 1. \tag{5}$$

(vi) *If there exists an intransitive Abelian group G′ containing G and having for one of its systems of intransitivity the sum of two or more systems of intransitivity of G, then the set of functions corresponding to this system of intransitivity of G′ does not satisfy all the conditions of Theorem* A.

Proof. The necessity of conditions (i), (ii), (iii), (vi) follows directly from Theorem A; the necessity of condition (iv) follows from the fundamental uniqueness theorem in the theory of almost-periodic functions; and the necessity of condition (v) may be shown by a slight modification of the proof of Theorem A.

We prove now the sufficiency of these conditions. Theorem A assures us that, for fixed k, the functions $f_{kh}(t)$ $(h = 1, \cdots, m_k)$ are distinct. Condition (iv) tells us that no two functions drawn from different systems of intransitivity are identical, so that all the functions must be distinct. Now $[f(t)]$ must have a certain Abelian almost-translation group G'.[4] Again using a modification of the

[4] Loc. cit., p. 11, Theorem 1.

proof of Theorem A, we can show that condition (v) compels every substitution in G to be in G', and excludes from G' every substitution on the systems of intransitivity of G which is not in G. Condition (vi) prevents G' from containing any system of intransitivity which is composed of two or more systems of intransitivity of G. Hence we can assert that $G' \equiv G$, and Theorem 1 is completely proved.

For our later work we replace condition (vi) by a sufficient but not necessary one by using Theorem A:

(vi)′ $|\, a_{ki} \,| \neq |\, a_{hj} \,|$, $\quad k \neq h$, *for every* i, j.

We proceed now to develop our most important result, that for an arbitrary intransitive Abelian group G we can always find a set of distinct almost-periodic functions which has G as its almost-translation group. Suppose G is of order N and involves r systems of intransitivity. It can be expressed as a direct product of cyclic subgroups, and generators of these cyclic subgroups form a minimal basis for G. Suppose that $S(1), \cdots, S(M)$ form such a minimal basis; $S(j)$ has order $N(j) \leqq N$. For each j $(1 \leqq j \leqq M)$ we write

$$S(j) = \prod_k s(k, j) \qquad (k = 1, \cdots, r),$$

where $s(k, j)$ belongs to the transitive group $G(k)$, and where for fixed k, the not necessarily distinct substitutions $s(k, j)$ generate $G(k)$. Selecting a minimal basis of $G(k)$ which contains, let us say, w_k elements, with $w_k \leqq M$, we write the members of this minimal basis of $G(k)$ in the form $T(k, h)$ $(h = 1, \cdots, w_k)$. We denote the order of $T(k, h)$ by $N(k, h)$, with $N(k, h) \leqq N$. For any $s(k, j)$ $(1 \leqq j \leqq M)$ we can find a set of integral c's for which

$$s(k, j) = \prod_h [T(k, h)]^{c_{kjh}}.$$

Finally, we rewrite the generators of G in the form

$$S(j) = \prod_{k,h} [T(k, h)]^{c_{kjh}}$$

$$(j = 1, \cdots, M; k = 1, \cdots, r; h = 1, \cdots, w_k).$$

If q' is the greatest w_k and $q = \sum w_k$, then $q' \leqq M \leqq q$. If $M = q$, we say that we have the *maximal group* G on the given systems of intransitivity. First assume that $M < q$. From the given group G we can always form the maximal group by adding, say $M' - M = p$ $(M' \geqq q)$, additional generating elements: $S(M + 1), \cdots, S(M')$.

We start our construction by forming the character-group $G^*(k)$ isomorphic to $G(k)$. We seek a basis of $G^*(k)$ containing w_k elements which are images of the $T(k, h)$. The following definition of characters over just the basis of $G(k)$ furnishes us with a desired basis for $G^*(k)$ and will satisfy condition (iii).

$$\chi_{kl}[T(k, h)] = \begin{cases} \exp\,(i2\pi/N(k, h)), & l = h, \\ 1, & l \neq h, \end{cases} \tag{8}$$

where $l, h = 1, \cdots, w_k$.

Corresponding to each generator $S(j) = \prod_{k,h} [T(k, h)]^{c_{kjh}}$ of G, we form the product

$$\prod_{k,l,h} \chi_{kl}\{[T(k, h)]^{c_{kjh}}\}$$

which according to (8) becomes

$$\prod_{k,h} \exp\,(i2\pi c_{kjh} d_{kh} g_{kh}/N),$$

with $d_{kh} = N/N(k, h)$. Whether or not this product is one depends on whether or not

$$\sum_{k,h} c_{kjh} d_{kh} g_{kh} \equiv 0 \pmod N \qquad (j = 1, \cdots, M). \tag{9}$$

For any substitution in the maximal group on the systems of intransitivity of G we can arrive at a similar congruence

$$\sum_{k,h} c_{kh} d_{kh} g_{kh} \equiv 0 \pmod N. \tag{10}$$

Consider a congruence (10) which corresponds to a substitution of G. There must exist a set of integral u's such that $S = \prod_j [S(j)]^{u_j}$ or such that

$$c_{kh} \equiv \sum_j c_{kjh} u_j \pmod N \qquad (j = 1, \cdots, M) \tag{11}$$

for every $k = 1, \cdots, r$; $h = 1, \cdots, w_k$. For any substitution generated by $S(M + 1), \cdots, S(M')$ there does not exist a set of integral u's satisfying (11).

Because of (11), every solution of (9) also satisfies every congruence (10) corresponding to a substitution of G, and hence satisfies (3) for every substitution of G. For every set of c_{kh}'s in (10) corresponding to a substitution generated by $S(M + 1), \cdots, S(M')$, we select a set of relatively prime g's which satisfies (9) but not (10),[5] i.e., which satisfies (3) for every substitution of G and (5) for this substitution not in G. Since $S(M + 1), \cdots, S(M')$ generate p elements, we shall obtain p sets of g's satisfying (9) but not (10) for at least one of these p elements. We write these sets

$$\{g_{khj}, k = 1, \cdots, r; h = 1, \cdots, w_k\} \qquad (j = 1, \cdots, p).$$

For every j $(1 \leqq j \leqq p)$ consider the equation

$$\sum_{k,h} g_{khj} \lambda_{khj} = 0, \qquad \sum w_k = q. \tag{12}$$

If we take $q - 1$ of the λ's as rationally independent, the remaining λ will be determined from (12). This choice and the fact that the g's are relatively prime

[5] It can be shown by standard methods from the theory of numbers that, given a system of congruences (9) and an additional congruence (10), the system (9) always has a relatively prime set of solutions which does not satisfy (10) if the left member of (10) is not identically congruent mod N to a linear integral combination of the left members of (9).

will prevent the λ's from satisfying any other relation of the form $\sum_{k,h} g'_{khj}\lambda_{khj} = 0$ except if the g''s are constant integral multiples of the corresponding g's. If the above sets of λ's are not rationally independent, then by multiplying as many as necessary of these p sets by suitable irrational constant factors, we can obtain p rationally independent sets of λ's. To each set of λ's we assign characters as defined in (8). There will be p sets

$$\{\chi_{khj}, h = 1, \cdots, w_k; k = 1, \cdots, r\} \qquad (j = 1, \cdots, p)$$

such that, for fixed k and h, all the χ_{khj} are equal.

This choice of the λ's and characters, as we show now, completely satisfies condition (v) of Theorem 1. The only relations, to within constant integral multiples, which the λ's satisfy are those in (12). For each such relation, (2) implies (3) for every substitution in G; for every substitution generated by $S(M+1), \cdots, S(M')$, there exists a relation (12) such that (4) and (5) hold true. Therefore condition (v) is disposed of. Furthermore, the characters have been chosen in accordance with condition (iii).

If $M = q$, we merely select one set of q exponents, all rationally independent, and one set of characters according to (8). In this case condition (v) is still satisfied, since the second part of the condition drops away (now every substitution on the systems of intransitivity of G is in G, so that $p = 0$).

In either case ($M < q$ or $M = q$), for each exponent we choose a coefficient different from zero such that all the coefficients obtained differ from each other in absolute value. This choice will satisfy conditions (iv) and (vi)$'$. Finally, we set up the functions

$$f_{kh} = \sum_{j,n} \chi_{knj}[S(k, h)]\, a_{knj} \exp(i\lambda_{knj}t) \qquad (h = 1, \cdots, m_k; k = 1, \cdots, r)$$

if $M < q$; and if $M = q$, the functions

$$f_{kh} = \sum_{n} \chi_{kn}[S(k, h)]\, a_{kn} \exp(i\lambda_{kn}t).$$

In these representations we remember that the substitution $S(k, h)$ takes f_{k1} into f_{kh}. The above functions in either case satisfy conditions (i) and (ii), and since they already have been shown to satisfy conditions (iii), (iv), (v), and (vi)$'$, we have the final

THEOREM 2. *Corresponding to each arbitrarily assigned intransitive Abelian group G having r systems of intransitivity with m_k integers in the k-th one, there exists a set $[f(t)]$ of $\sum m_k$ distinct almost-periodic functions which has G for its almost-translation group.*

We conclude this paper with a list of examples. Taking the integers 1, 2, 3, 4, 5, 6 with the systems of intransitivity 1, 2; 3, 4; 5, 6, we form every possible intransitive Abelian group G on these systems and then construct corresponding sets of distinct almost-periodic functions having G for their almost-translation group. For the substitutions on the systems of intransitivity we write

$S_1 = (12)$, $S_2 = (34)$, $S_3 = (56)$, and for the corresponding characters we write $\chi_1(S_1) = \chi_2(S_2) = \chi_3(S_3) = -1$. a_1, a_2, a_3 will denote complex constants different from zero and from each other in absolute value; λ, μ, ν will denote rationally independent real numbers. With this understanding we have the following possibilities for the group G:

(a) *G, of order 2, generated by $S_1S_2S_3$.* As solutions of the congruence $g_1 + g_2 + g_3 \equiv 0 \pmod 2$, we take $g_1 = 1$, $g_2 = -1$, $g_3 = 0$ and $g_1 = 0$, $g_2 = 1$, $g_3 = -1$. From $\lambda_1 - \lambda_2 = 0$ and $\lambda_2 - \lambda_3 = 0$ take $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$. We construct the functions

$$f_1 = a_1 \exp(i\lambda t), \quad f_3 = a_2 \exp(i\lambda t), \quad f_5 = a_3 \exp(i\lambda t),$$
$$f_2 = -a_1 \exp(i\lambda t), \quad f_4 = -a_2 \exp(i\lambda t), \quad f_6 = -a_3 \exp(i\lambda t).$$

(b) *G, of order 4, generated by S_1S_2 and S_3.* For a solution of $g_1 + g_2 \equiv 0 \pmod 2$, and $g_3 \equiv 0 \pmod 2$, take $g_1 = 1$, $g_2 = -1$, $g_3 = 0$, giving $\lambda_1 - \lambda_2 + 0\lambda_3 = 0$. Choose $\lambda_1 = \lambda_2 = \lambda$, $\lambda_3 = \mu$. For the functions we take

$$f_1 = a_1 \exp(i\lambda t), \quad f_3 = a_2 \exp(i\lambda t), \quad f_5 = a_3 \exp(i\mu t),$$
$$f_2 = -a_1 \exp(i\lambda t), \quad f_4 = -a_2 \exp(i\lambda t), \quad f_6 = -a_3 \exp(i\mu t).$$

(c) *G, of order 4, generated by S_1S_2 and S_1S_3.* For a solution of $g_1 + g_2 \equiv 0 \pmod 2$, and $g_1 + g_3 \equiv 0 \pmod 2$, take $g_1 = 1$, $g_2 = g_3 = -1$, giving $\lambda_1 - \lambda_2 - \lambda_3 = 0$. Take $\lambda_1 = \lambda + \mu$, $\lambda_2 = \lambda$, $\lambda_3 = \mu$, and for the functions use

$$f_1 = a_1 \exp[i(\lambda + \mu)t], \quad f_3 = a_2 \exp(i\lambda t), \quad f_5 = a_3 \exp(i\mu t),$$
$$f_2 = -a_1 \exp[i(\lambda + \mu)t], \quad f_4 = -a_2 \exp(i\lambda t), \quad f_6 = -a_3 \exp(i\mu t).$$

(d) *G, of order 8, generated by S_1, S_2, and S_3.* Take the functions

$$f_1 = a_1 \exp(i\lambda t), \quad f_3 = a_2 \exp(i\mu t), \quad f_5 = a_3 \exp(i\nu t),$$
$$f_2 = -a_1 \exp(i\lambda t), \quad f_4 = -a_2 \exp(i\mu t), \quad f_6 = -a_3 \exp(i\nu t).$$

Any other intransitive Abelian group G on 1, 2; 3, 4; 5, 6 after a proper rearrangement of these integers will fall into type (b).

NEW YORK UNIVERSITY.

RIEMANN SUMS AND THE FUNDAMENTAL POLYNOMIALS OF LAGRANGE INTERPOLATION

By J. H. Curtiss

1. **Introduction.** Let C denote an arbitrary Jordan curve of the complex z-plane, and let $z = \varphi(w)$ be an analytic function which maps the exterior K of C (i.e., the unlimited region bounded by C) conformally onto the region $|w| > 1$ so that the points at infinity correspond. We assume that this function is defined so as to be continuous and univalent for $1 \leqq |w| < \infty$. The Laurent series for the function can be written as follows:

$$\varphi(w) = cw + c_0 + \frac{c_1}{w} + \frac{c_2}{w^2} + \cdots, \qquad |w| \geqq 1, \tag{1.1}$$

where $|c|$ is the transfinite diameter of C.

The polynomials

$$\omega_n(z) = \prod_{k=1}^{n} [z - \varphi(e^{2\pi ik/n})] \qquad (n = 1, 2, \cdots)$$

are called the fundamental polynomials of Lagrange interpolation in the points $\varphi(e^{2\pi ik/n})$ on C. It is well known[1] that

$$\lim_{n\to\infty} |\omega_n(z)|^{1/n} = \begin{cases} |c|\,|w|, & z = \varphi(w),\ z \text{ in } K, \\ |c|, & z \text{ interior to } C. \end{cases}$$

In the present paper we attack the more delicate problem of determining the exact behavior of the sequence $\{\omega_n(z)\}$, rather than that of the sequence $\{|\omega_n(z)|^{1/n}\}$. The results to be established may be stated formally as follows:

THEOREM 1. *If C is rectifiable, then*

$$\lim_{n\to\infty} \frac{\omega_n(z)}{-c^n} = 1 \tag{1.2}$$

uniformly for z on any closed set M of the region interior to C, and

$$\lim_{n\to\infty} \frac{\omega_n(z)}{c^n(w^n - 1)} = 1, \qquad z = \varphi(w), \tag{1.3}$$

uniformly for z on any closed set M_1 of K.

Received March 22, 1941.

[1] L. Fejér, *Interpolation und konforme Abbildung*, Göttinger Nachrichten, 1918, pp. 319–331; J. L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, American Mathematical Society Colloquium Publications, vol. 20, New York, 1935; especially pp. 68–75.

THEOREM 2. *If $\varphi'(w)$ is non-vanishing and of bounded variation for $|w| = 1$, then (1.3) is true uniformly for z in the closed region $C + K$.*

(By $\varphi'(w)$ we mean the function $c - c_1w^{-2} - 2c_2w^{-3} - \cdots$.)

We remark that an interesting limiting case is that in which C reduces to the line segment $-1 \leqq z \leqq 1$, and $\varphi(w) = \frac{1}{2}(w + 1/w)$, $\varphi(e^{2\pi ik/n}) = \cos(2\pi k/n)$ $(k = 1, \cdots, n;\ n = 1, 2, \cdots)$. This set of points includes the abscissas of Tchebycheff. Here $\omega_n(z)$ is easily found by direct computation, and turns out to be $2^{-n}(w^n - 1)(1 - w^{-n})$. Hence (1.3) is satisfied in K but not on C.

Theorems 1 and 2 were proved earlier by the author[2] under considerably more restrictive conditions on C, which essentially involved Lipschitz conditions on the derivatives of $\varphi(w)$. Recently Walsh and Sewell[3] have obtained bounds for the sequence $\{|\omega_n(z)|\}$ in the case of several curves, using restrictions on the curves which involve a Lipschitz condition imposed on the logarithm of the difference-quotient of the mapping function. We have two purposes in re-opening here the discussion of the case of one curve. The first purpose is to establish Theorems 1 and 2 under lighter and possibly more "natural" conditions than those hitherto used in such problems. To this end we shall employ a method of proof suggested by the work of Walsh and Sewell, in which the logarithm of $\omega_n(z)/(-c^n)$, or of $\omega_n(z)/[c^n(w^n - 1)]$, is treated as a Riemann sum. The second purpose is to present auxiliary results, which are perhaps of some independent interest, on the degree of convergence of the Riemann sums for absolutely continuous functions. These auxiliary results are contained in the next section.

2. Degree of convergence of certain Riemann sums.

LEMMA 1. *Let $f(\theta)$ be a complex function of the real variable θ absolutely continuous in the interval $a \leqq \theta \leqq b$. Let $\theta_k = \theta_k^{(n)} = a + (k/n)(b - a)$ $(k = 0, 1, 2, \cdots, n)$. Then*

$$\lim_{n\to\infty} n\left[\sum_{k=1}^{n} f(\theta_k)(\theta_k - \theta_{k-1}) - \int_a^b f(\theta)\,d\theta\right] = \frac{b-a}{2}[f(b) - f(a)]. \tag{2.1}$$

This result is already known in the special case in which $df/d\theta$ is bounded and integrable in the sense of Riemann.[4] A proof of the more general case can be given along the following lines:

Integrating by parts, we find that

$$\left[\sum_{k=1}^{n} f(\theta_k)(\theta_k - \theta_{k-1}) - \int_a^b f(\theta)\,d\theta\right] = \sum_{k=1}^{n} \int_{\theta_{k-1}}^{\theta_k} (\theta - \theta_{k-1})\left(\frac{df}{d\theta}\right) d\theta. \tag{2.2}$$

[2] J. H. Curtiss, *Interpolation in regularly distributed points*, Transactions of the American Mathematical Society, vol. 38(1935), pp. 458–473; §4. As a result of the work in the present paper, the restriction that the curve C should satisfy "condition (a)" in Theorem I of the Transactions paper may be replaced by the restriction that C be merely rectifiable, and the conditions on C in the hypotheses of Theorems II and III of that paper may be replaced by the conditions in the hypothesis of Theorem 2 above.

[3] J. L. Walsh and W. E. Sewell, *Sufficient conditions for various degrees of approximation by polynomials*, this Journal, vol. 6(1940), pp. 658–705.

[4] See G. Pólya and G. Szegö, *Aufgaben und Lehrsätze*, Berlin, 1925; vol. I, p. 37.

Letting

$$D_n = \sum_{k=1}^{n} \int_{\theta_{k-1}}^{\theta_k} \left(\frac{df}{d\theta}\right)\left[n(\theta - \theta_{k-1}) - \frac{b-a}{2}\right] d\theta,$$

we see that (2.1) is equivalent to the assertion that $\lim_{n\to\infty} D_n = 0$. The proof of this latter relation is very similar to the standard proof of the Riemann-Lebesgue theorem in the theory of Fourier series.[5] Let ϵ be an arbitrary positive number, and let $g(\theta)$ be a function absolutely continuous in the interval $a \leqq \theta \leqq b$ and such that

$$\int_a^b \left|\frac{df}{d\theta} - g(\theta)\right| d\theta < \frac{\epsilon}{(b-a)}.$$

We define d_n by the equation

$$d_n = \sum_{k=1}^{n} \int_{\theta_{k-1}}^{\theta_k} g(\theta)\left[n(\theta - \theta_{k-1}) - \frac{b-a}{2}\right] d\theta$$
$$= \sum_{k=1}^{n} \int_{\theta_{k-1}}^{\theta_k} \left(\frac{dg}{d\theta}\right)\left[\frac{(\theta - \theta_{k-1})(b-a)}{2} - n\frac{(\theta - \theta_{k-1})^2}{2}\right] d\theta.$$

If θ lies in the interval $\theta_{k-1} \leqq \theta \leqq \theta_k$, then $0 \leqq \theta - \theta_{k-1} \leqq (b-a)/n$, so

$$|d_n| \leqq \frac{(b-a)^2}{2n} \int_a^b \left(\frac{dg}{d\theta}\right) d\theta.$$

We then have the inequality

$$(2.3)\quad \begin{aligned} |D_n| \leqq |D_n - d_n| + |d_n| &\leqq \frac{b-a}{2} \int_a^b \left|\frac{df}{d\theta} - g(\theta)\right| d\theta \\ &+ \frac{(b-a)^2}{2n} \int_a^b \left|\frac{dg}{d\theta}\right| d\theta \leqq \frac{\epsilon}{2} + \frac{\epsilon}{2}, \qquad n > N_\epsilon, \end{aligned}$$

and this concludes the proof.

The lemma shows that the first member of (2.2) is $O(1/n)$ if $f(b) \neq f(a)$ and is $o(1/n)$ if $f(b) = f(a)$.

LEMMA 2. *Let $f(\theta, \alpha)$ be a complex function of the real variable θ and the complex variable α which satisfies the following conditions:*

[5] See, for instance, E. C. Titchmarsh, *The Theory of Functions*, Oxford, 1932; pp. 403–404. This similarity was pointed out by the referee, who remarked that the relation $\lim_{n\to\infty} D_n = 0$ was a special case of the following general Riemann-Lebesgue theorem:

Let $F(x)$ be a bounded, periodic, measurable function over $-\infty < x < \infty$ whose integral over a period vanishes. Let $f(x)$ be any integrable function on $0 \leqq x \leqq 1$. Then $\lim_{y\to\infty} \int_0^1 f(x)F(xy)\, dx = 0$.

Our chief interest in the present paper lies in an extension of Lemma 1 to the case in which $f(\theta)$ depends upon a parameter α; this extension appears in Lemma 2. Although the remaining details of the proof of Lemma 1 could easily be supplied by the reader, it seems worth while to arrange them here in such a way that the argument will apply at once to the extended form of the result.

(A) *For each value of α on a set S, $f(\theta, \alpha)$ is an absolutely continuous function of θ in the interval $a \leqq \theta \leqq b$.*

(B) *For any given $\epsilon > 0$, there exists a function $g(\theta, \alpha)$ absolutely continuous in θ in the interval $a \leqq \theta \leqq b$, for each α on S, and such that*

$$\int_a^b \left| \frac{\partial f}{\partial \theta} - g(\theta, \alpha) \right| d\theta < \epsilon, \qquad \alpha \in S, \tag{2.4}$$

$$\int_a^b \left| \frac{\partial g}{\partial \theta} \right| d\theta < M(\epsilon), \qquad \alpha \in S, \tag{2.5}$$

where $M(\epsilon)$ is independent of α.

Then

$$\lim_{n \to \infty} n \left[\sum_{k=1}^{n} f(\theta_k, \alpha)(\theta_k - \theta_{k-1}) - \int_a^b f(\theta, \alpha)\, d\theta \right]$$

$$= \frac{b-a}{2} [f(b, \alpha) - f(a, \alpha)], \quad \textit{uniformly}, \qquad \alpha \in S,$$

where $\theta_k = a + (k/n)(b - a)$.

The proof is essentially the same as that of Lemma 1; a glance at (2.3) will indicate where and how (2.4) and (2.5) are to be used.

3. **Proof of Theorems 1 and 2.** We first establish (1.2). The function $[\varphi(\bar{w}) - z]/c\bar{w}$, $z \in M$, is analytic in the variable $\bar{w}$ and non-vanishing for $|\bar{w}| > 1$, and is continuous for $|\bar{w}| \geqq 1$ and equal to one at infinity when defined there by continuity. Let $\log \{[\varphi(\bar{w}) - z]/c\bar{w}\}$ denote a branch of the logarithm of this function chosen so that it is single valued and analytic for $|\bar{w}| > 1$, continuous for $|\bar{w}| > 1$, and equal to zero at infinity. We write

$$F(\theta, z) = \log \left[\frac{\varphi(e^{i\theta}) - z}{ce^{i\theta}} \right], \qquad z \in M. \tag{3.1}$$

By Cauchy's Integral Theorem for the closed region $|\bar{w}| \geqq 1$,

$$\int_0^{2\pi} F(\theta, z)\, d\theta = \int_{|\bar{w}|=1} \log \left[\frac{\varphi(\bar{w}) - z}{c\bar{w}} \right] \frac{d\bar{w}}{i\bar{w}} = 0.$$

Now

$$\frac{\omega_n(z)}{-c^n} = -\prod_{k=1}^{n} \left[\frac{z - \varphi(e^{i\theta_k})}{c} \right] = \prod_{k=1}^{n} \left[\frac{\varphi(e^{i\theta_k}) - z}{ce^{i\theta_k}} \right],$$

where $\theta_k = 2\pi k/n = 0 + (k/n)(2\pi - 0)$. A properly chosen branch of the logarithm of $\omega_n(z)/(-c^n)$ will satisfy the equation

$$2\pi \log \frac{\omega_n(z)}{-c^n} = 2\pi \sum_{k=1}^{n} F(\theta_k, z) = n \sum_{k=1}^{n} F(\theta_k, z)(\theta_k - \theta_{k-1})$$

$$= n \left[\sum_{k=1}^{n} F(\theta_k, z)(\theta_k - \theta_{k-1}) - \int_0^{2\pi} F(\theta, z)\, d\theta \right].$$

Thus, according to Lemma 2, we shall have established that

$$\lim_{n\to\infty} 2\pi \log \frac{\omega_n(z)}{-c^n} = \frac{2\pi - 0}{2}[F(2\pi, z) - F(0, z)] = 0, \quad \text{uniformly}, \qquad z \in M$$

(which is the same as (1.2)), when we have proved the following assertion:

LEMMA 3. *The function* $F(\theta, z)$ *satisfies conditions* (A) *and* (B) *of Lemma* 2 *with* α *replaced by* z, S *by* M, a *by* 0, b *by* 2π.

Proof. (i) F *satisfies* (A). For if C is rectifiable, then $\varphi(e^{i\theta})$ is of bounded variation, $0 \leqq \theta \leqq 2\pi$, and therefore the unit circle boundary values of the function $w'\varphi(1/w')$, which is analytic for $|w'| < 1$, are of bounded variation. By a theorem of F. and M. Riesz,[6] it follows that $e^{-i\theta}\varphi(e^{i\theta})$ must be absolutely continuous, $0 \leqq \theta \leqq 2\pi$. Accordingly, F is absolutely continuous in θ ($0 \leqq \theta \leqq 2\pi$) for each $z \in M$.

(ii) F *satisfies* (B). We have

$$\frac{\partial F}{\partial \theta} = \frac{1}{\varphi(e^{i\theta}) - z}\frac{d\varphi}{d\theta} - i.$$

Given any $\epsilon > 0$, let $h(\theta)$ be an absolutely continuous function such that

$$\int_0^{2\pi} \left| \frac{d\varphi}{d\theta} - h(\theta) \right| d\theta < \epsilon d,$$

where d is the distance between M and C. Construct the function

$$G(\theta, z) = \frac{h(\theta)}{\varphi(e^{i\theta}) - z} - i.$$

It is now a trivial exercise to show that (2.4) and (2.5) are satisfied by G and F. The proof of Lemma 3 is complete.

We turn to (1.3) and Theorem 2. The remarks made at the beginning of this section concerning the function $[\varphi(\bar{w}) - z]/c\bar{w}$ apply to the function

$$q(\bar{w}, w) = \begin{cases} \dfrac{\varphi(\bar{w}) - \varphi(w)}{c(\bar{w} - w)}, & \bar{w} \neq w, \\ \varphi'(w), & \bar{w} = w, \end{cases}$$

with the understanding that if the hypothesis of Theorem 1 is in effect, the fixed point w is to be taken in the region $|w| > 1$, and if the hypothesis of Theorem 2 is in effect, w is to lie in the closed region $|w| \geqq 1$. In either case, let $\psi(\theta, w) = \log q(e^{i\theta}, w)$, where the branch of the logarithm is chosen as in (3.1). Then again

$$\int_0^{2\pi} \psi(\theta, w)\, d\theta = 0,$$

[6] F. and M. Riesz, *Über Randwerte einer analytischen Funktion*, Comptes Rendus du Quatrième Congrès (1916) des Mathématiciens Scandinaves, Uppsala, 1920, pp. 27–44. See also A. Zygmund, *Trigonometrical Series*, Warsaw, 1935, pp. 158–162.

and

$$2\pi \log \frac{\omega_n(z)}{c^n(w^n - 1)} = n\left[\sum_{k=1}^{n} \psi(\theta_k, w)(\theta_k - \theta_{k-1}) - \int_0^{2\pi} \psi(\theta, w)\, d\theta\right].$$

Thus according to Lemma 2, we shall have proved that

$$\lim_{n\to\infty} 2\pi \log \frac{\omega_n(z)}{c^n(w^n - 1)} = \frac{2\pi - 0}{2}[\psi(2\pi, w) - \psi(0, w)] = 0, \qquad z = \varphi(w),$$

uniformly for z on M_1 (Theorem 1) or z on C (Theorem 2), when the following results have been established:

LEMMA 4. *If C is rectifiable, then $\psi(\theta, w)$ satisfies conditions* (A) *and* (B) *with α replaced by w, S by the image in the w-plane of M_1, a by 0, b by 2π.*

LEMMA 5. *If $\varphi'(w)$ is non-vanishing and of bounded variation for $|w| = 1$, then $\psi(\theta, w)$ satisfies conditions* (A) *and* (B) *with α replaced by w, S by $|w| = 1$, a by 0, b by 2π.*

Of course, if (1.3) holds uniformly for z on C, then it holds uniformly for z on $C + K$, by the Principle of the Maximum.

The proof of Lemma 4 is so similar to that of Lemma 3 that it does not seem necessary to present it here. But Lemma 5 lies somewhat deeper, and since the result is apparently of interest apart from the context, we shall occupy the remainder of this paper with a sketch of the proof.

We first prove Lemma 5 for the difference-quotient $q(e^{i\theta}, w)$ instead of for $\psi(\theta, w)$. The restriction that $\varphi'(w) \neq 0$, $|w| = 1$, is here superfluous.

(i) *$q(e^{i\theta}, w)$ satisfies* (A). Let $\bar{w} = e^{i\theta}$, $w = e^{i\beta}$, and $u = e^{it}$. By the theorem of F. and M. Riesz referred to above, $\varphi'(w)$ is absolutely continuous for $|w| = 1$. The periodic function $\partial q/\partial\theta$ is clearly continuous in θ except possibly for $\theta \equiv \beta \pmod{2\pi}$. Thus a sufficient condition for the absolute continuity of q over any finite interval of the θ-axis for each fixed value of w (or β) is that

$$\left(\int_{\beta-\pi}^{\beta-\eta} + \int_{\beta+\eta'}^{\beta+\pi}\right)\left|\frac{\partial q}{\partial\theta}\right| d\theta$$

remain bounded as η and $\eta' \to 0$, where $\eta > 0$ and $\eta' > 0$.[7] Assume that $\eta < \pi$, $\eta' < \pi$. Integration by parts yields

$$\text{(3.2)} \qquad \frac{\partial q}{\partial\theta} = \frac{\partial q}{\partial\bar{w}} \cdot \frac{d\bar{w}}{d\theta} = \frac{i\bar{w}}{c} \frac{\int_\beta^\theta (u - w)\varphi''(u) iu\, dt}{(\bar{w} - w)^2}, \qquad \bar{w} \neq w.$$

Then since $|\bar{w} - w| = 2|\sin \tfrac{1}{2}(\theta - \beta)|$,

$$\begin{aligned} I &= |c|\left(\int_{\beta-\pi}^{\beta-\eta} + \int_{\beta+\eta'}^{\beta+\pi}\right)\left|\frac{\partial q}{\partial\theta}\right| d\theta \\ &\leqq \left(\int_{\beta-\pi}^{\beta-\eta} + \int_{\beta+\eta'}^{\beta+\pi}\right) \tfrac{1}{2}\csc^2 \tfrac{1}{2}(\theta - \beta) \int_\beta^\theta \sin \tfrac{1}{2}(t - \beta)\, |\varphi''(u)|\, dt\, d\theta. \end{aligned}$$

[7] See, for instance, E. C. Titchmarsh, *The Theory of Functions*, Oxford, 1932; p. 372, ex. 6.

We integrate the right member by parts, and obtain

$$
\begin{aligned}
I \leqq \operatorname{ctn} \tfrac{1}{2}\eta \int_{\beta}^{\beta-\eta} & \sin \tfrac{1}{2}(t-\beta)\,|\varphi''(u)|\,dt \\
+ \operatorname{ctn} \tfrac{1}{2}\eta' \int_{\beta}^{\beta+\eta'} & \sin \tfrac{1}{2}(t-\beta)\,|\varphi''(u)|\,dt \\
+ \left(\int_{\beta-\pi}^{\beta-\eta} + \int_{\beta+\eta'}^{\beta+\pi}\right) & \cos \tfrac{1}{2}(\theta-\beta)\,|\varphi''(\bar{w})|\,d\theta \leqq 3\int_0^{2\pi} |\varphi''(\bar{w})|\,d\theta.
\end{aligned} \tag{3.3}
$$

Thus I remains bounded as η and $\eta' \to 0$. We have shown that q satisfies (A). A slight extension of this argument, which makes use of the fact that q is symmetric in θ and β, serves to show that q is a uniformly continuous function of the *two* variables θ and β in the closed square $0 \leqq \theta \leqq 2\pi$, $0 \leqq \beta \leqq 2\pi$. Hence if $\varphi'(w) \neq 0$, $|w| = 1$, there exists a positive number m such that $|q| \geqq m$ for all θ and β—a fact which we shall presently use in studying $\psi(\theta, w)$.

(ii) $q(e^{i\theta}, w)$ *satisfies* (B). From (1.1), we find that

$$\varphi''(\bar{w}) = \frac{2c_1}{\bar{w}^3} + \frac{2\cdot 3c_2}{\bar{w}^4} + \frac{3\cdot 4c_3}{\bar{w}^5} + \cdots, \qquad |\bar{w}| > 1; \tag{3.4}$$

and it is known that this Laurent series becomes the Fourier series generated by $\varphi''(e^{i\theta})$ if, as henceforth, we set $\bar{w} = e^{i\theta}$.[8] Let $\{\sigma_N(\bar{w})\}$ denote the sequence of arithmetic means of this series, and let $\epsilon_N = \int_0^{2\pi} |\varphi''(\bar{w}) - \sigma_N(\bar{w})|\,d\theta$. Then by a theorem of Steinhaus,[9] $\lim_{n\to\infty} \epsilon_N = 0$. We now introduce the function

$$\nu_N(\bar{w}, w) = \begin{cases} \dfrac{\dfrac{i\bar{w}}{c}\displaystyle\int_{\beta}^{\theta} (u-w)\sigma_N(u)iu\,dt}{(\bar{w}-w)^2}, & \bar{w} \neq w, \\[2ex] \dfrac{iw}{2c}\sigma_N(w), & \bar{w} = w, \end{cases}$$

where as usual $w = e^{i\beta}$, $u = e^{it}$. The reader can verify that the second quantity in the brace here is the limiting value of the first quantity as $\bar{w} \to w$. It is apparent from (3.4) that $\sigma_N(u)$ contains no power of $1/u$ lower than the third. Also,

$$\frac{\displaystyle\int_{\beta}^{\theta} (u-w)\frac{h(h+1)}{u^{h+2}}iu\,dt}{(\bar{w}-w)^2} = \frac{d}{d\bar{w}}\left[\frac{\bar{w}^{-h} - w^{-h}}{\bar{w}-w}\right] \qquad (h = 1, 2, \cdots)$$

(compare (3.2)), and this expression is simply a polynomial in $1/\bar{w}$ and $1/w$. Then $\nu_N(\bar{w}, w)$, being a linear combination of such polynomials, is itself a poly-

[8] F. and M. Riesz, loc. cit.; also Zygmund, op. cit., p. 158.

[9] H. Steinhaus, *Sur quelques propriétés des séries trigonométriques et de celles de Fourier*, Rozprawj Akademji Umiejętności, Cracow, vol. 58(1925), pp. 175–225; also Zygmund, op. cit., pp. 84–85.

nomial in $1/\bar{w}$ and $1/w$. From this it follows that ν_N is absolutely continuous in θ for each w, and that ν_N satisfies (2.5) for $|w| = 1$. Now

$$
\begin{aligned}
J_N &= |c| \int_0^{2\pi} \left| \frac{\partial q}{\partial \theta} - \nu_N(\bar{w}, w) \right| d\theta \\
&= |c| \lim_{\substack{\eta \to 0 \\ \eta' \to 0}} \left(\int_{\beta-\pi}^{\beta-\eta} + \int_{\beta+\eta'}^{\beta+\pi} \right) \left| \frac{\partial q}{\partial \theta} - \nu_N(\bar{w}, w) \right| d\theta \\
&\leqq \lim_{\substack{\eta \to 0 \\ \eta' \to 0}} \left(\int_{\beta-\pi}^{\beta-\eta} + \int_{\beta+\eta'}^{\beta+\pi} \right) \tfrac{1}{2} \csc^2 \tfrac{1}{2}(\theta - \beta) \int_\beta^\theta \sin \tfrac{1}{2}(t - \beta) \, |\varphi''(u) - \sigma_N(u)| \, dt \, d\theta, \\
&\qquad\qquad 0 < \eta < \pi, \; 0 < \eta' < \pi.
\end{aligned}
$$

Exactly the same methods used to derive (3.3) now yield

$$
J_N \leqq 3 \int_0^{2\pi} |\varphi''(\bar{w}) - \sigma_N(\bar{w})| \, d\theta = 3\epsilon_N.
$$

Since $\epsilon_N \to 0$ as $N \to \infty$, we have shown that q satisfies (B).

We turn now to $\psi(\theta, w) = \log q(\bar{w}, w)$. Using the facts that q is absolutely continuous and that $|q| \geqq m$, we can rapidly establish that ψ satisfies (A). Now $\partial\psi/\partial\theta = (1/q)(\partial q/\partial\theta)$, so as our approximating function we take $\Gamma_N(\theta, w) = (1/q)\nu_N(\bar{w}, w)$. Then

$$
\begin{aligned}
\int_0^{2\pi} \left| \frac{\partial \psi}{\partial \theta} - \Gamma_N(\theta, w) \right| d\theta &\leqq \frac{1}{m} \int_0^{2\pi} \left| \frac{\partial q}{\partial \theta} - \nu_N(\bar{w}, w) \right| d\theta \\
&= \frac{J_N}{m|c|} \leqq \frac{3\epsilon_N}{m|c|},
\end{aligned}
$$

which shows that the first member tends to zero as N becomes infinite. Thus for a suitably large value of N, ψ and Γ_N satisfy (2.4). It is easily seen that Γ_N satisfies (2.5) for a fixed value of N, so ψ satisfies (B), and the proof of Lemma 5 is complete.

Cornell University.

FINITE GROUPS AND RESTRICTED LIE ALGEBRAS

By Robert Hooke

1. **Introduction.** We are concerned with a method of Zassenhaus[1] which associates with abstract algebraic groups certain Lie rings.

By a Lie ring L over a field K, we mean a linear space (of finite or infinite dimension) over K, in which there is defined an operation $[x, y]$, called the commutator of the two elements x and y, satisfying

(a) $[x, y]$ is in L when x and y are in L.

(b) $[mx + ny, z] = m[x, z] + n[y, z]$, m, n in K, and a similar expression with the sum on the other side of the comma.

(c) $[x, x] = 0$ (and so $[x, y] = -[y, x]$).

(d) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$.

If L has a finite basis over K, we indicate this fact by calling L a Lie *algebra* over K.

If K is of characteristic p, we say that L is of characteristic p. An important class of Lie rings of characteristic p is the "restricted" Lie rings.[2] L is a restricted Lie ring if it has characteristic p, and if there is associated with every element x of L an element denoted by x^p which satisfies the condition

(e) $[y, x^p] = \overbrace{[\cdots [[y, x], x], \cdots, x]}^{p}$ for all y in L.

Let L and L' be two Lie rings over K with a one-to-one mapping $x \to x'$ of L onto L', such that $mx + ny \to mx' + ny'$ (m, n in K) and $[x, y] \to [x', y']$. Then L and L' are said to be isomorphic. If L and L' are restricted Lie rings and, in addition, $x^p \to (x')^p$, we shall say that L and L' are "p-isomorphic".

Let L be any Lie ring of characteristic p. We may or may not be able to choose for each element x an element x^p satisfying condition (e). If we are able to do so, we follow the nomenclature of Zassenhaus and call L a p-invariant Lie ring. For each element x there may exist several elements which would satisfy the condition (e). Indeed, let x' be such an element and C be the centrum (set of elements c such that $[c, y] = 0$ for all y in L); then $x' + c$ also satisfies (e) for any c in C.

Given a p-invariant Lie ring L, by choosing, for each element x, one element x^p which satisfies condition (e), L becomes a restricted Lie ring. Clearly we can begin with the same L and, by choosing different elements x^p, get restricted Lie rings which are not p-isomorphic. A methodical way of choosing the x^p is

Received April 24, 1941.

[1] H. Zassenhaus, *Ein Verfahren, jeder endlichen p-Gruppe einer Lie-Ring mit der Charakteristik p zuzuordnen*, Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität, vol. 13(1939), pp. 200–207.

[2] N. Jacobson, *Abstract derivation and Lie algebras*, Transactions of the American Mathematical Society, vol. 42(1937), pp. 206–224.

given by Zassenhaus. He shows that if L is p-invariant, the elements x^p may be so chosen as to satisfy the conditions:

(1) $(kx)^p = k^p x^p$, k in K.

(2) $(x + y)^p = x^p + y^p + S(x, y)$, where $S(x, y)$ is a linear combination of commutators of x and y.

He also shows that if this p-operation is to satisfy conditions (1) and (2), it is completely determined by its effect on a basis of L.

In the paper of Zassenhaus, there is described a method of associating with any group G, for any prime number p, a restricted Lie ring L over the Galois field $GF(p)$ of p elements; the p-operation induced by G in L satisfies conditions (1) and (2). We shall speak of L as the p-Lie ring of G. At the end of his paper, Zassenhaus states that he has not investigated the reverse process from L to G. The question here is: given a Lie ring L satisfying the obvious necessary conditions, is there a group G whose p-Lie ring is L? Or, more generally, is there a one-to-one correspondence between abstract groups and their p-Lie rings which is analogous to that between local Lie groups and their Lie algebras?

It will become obvious from the definitions that the p-Lie ring of any finite group is a Lie algebra and is nilpotent. It also happens that a p-group has only one non-zero Lie algebra, namely, the one which corresponds to the prime number which divides the order of the group. We shall show that if L is the p-Lie *algebra* of any group G, then it is the p-Lie algebra of a p-group G', and we shall show how G' arises from G. We might then hope to get a one-to-one correspondence between p-groups and restricted nilpotent Lie algebras, but examples will be given to show that:

(A) There exist non-isomorphic p-groups with Lie algebras which are p-isomorphic.

(B) There exists a nilpotent restricted Lie algebra L which arises from no p-group (and hence from no group whatsoever).

The statement (B) is to be read: There exists no group G whose p-Lie algebra is p-isomorphic to L. The weaker form of this statement, with "p-isomorphic" replaced by "isomorphic", is much less interesting, and remains unproved.

2. Summary of the results of Zassenhaus. We shall designate by Z_i the groups of the upper derived series (lower central series, absteigende Zentralreihe) of a group G. These are defined by

$$Z_1 = G, \quad Z_2 = [Z_1, G], \quad \cdots, \quad Z_n = [Z_{n-1}, G], \quad \cdots.$$

We use these subgroups of G to define another sequence of characteristic subgroups of G:

Definition. Given a group G and a prime number p, then G_n, the n-th dimension group mod p of G, is given by

$$G_n = \{Z_i^{p^j}\}, \qquad ip^j \geq n.$$

That is, G_n is the subgroup generated by all elements $z_i^{p^j}$, where z_i is any element of Z_i and i, j take all possible non-negative integral values such that $ip^j \geqq n$.

The groups G_n are shown to form a descending (not necessarily properly descending) chain of characteristic subgroups of G such that

$$G_1 = G, \quad [G_n, G_m] \subseteq G_{n+m}, \quad G_n^p \subseteq G_{np}.$$

The notation H^p for a subset H of G signifies the set of elements h^p for all elements h of H. The symbol $\{H^p\}$ will denote the subgroup generated by H^p.

From these results it is clear that the quotient groups G_n/G_{n+1} are all Abelian and of type $(p, p, \cdots)$ or 1. We can therefore map each G_n/G_{n+1} isomorphically onto a linear space $G_{(n)}$ over the field $GF(p)$ by a mapping which we shall call d_n.

The p-Lie ring of a group G can now be defined. Let L be the linear space over $GF(p)$ which is the direct sum of the $G_{(n)}$:

$$L = G_{(1)} + G_{(2)} + \cdots .$$

The non-zero elements of $G_{(m)}$ are $d_m g_m$, where g_m is in G_m but not in G_{m+1}. We define

$$[d_n g_n, d_m g_m] = d_{n+m}[g_n, g_m].$$

This can be extended uniquely and without contradiction to give a Lie multiplication for the whole of L. Finally, if we define

$$(d_n g_n)^p = d_{np} g_n^p,$$

it is shown that L is a restricted Lie ring over $GF(p)$, and the p-operation satisfies conditions (1), (2).

Let L be the p-Lie ring of a group G. We define the ideals of the upper derived series of L as

$$Z_1(L) = L, \quad Z_2(L) = [Z_1(L), L], \quad \cdots, \quad Z_n(L) = [Z_{n-1}(L), L], \quad \cdots .$$

Then if L_n and $L_{(n)}$ are defined as

$$L_n = \{Z_i(L)^{p^j}\}, \qquad ip^j \geqq n,$$

and

$$L_{(n)} = L_n/L_{n+1},$$

we have

$$L = L_{(1)} + L_{(2)} + \cdots ;$$

and this is the same subdivision of L into linear spaces that was given by the spaces $G_{(n)}$.

Finally, for a p-group, the dimension groups end at 1, and the p-Lie algebra has the same number of elements as the group. That is, if the order of a group is p^n, the order of its Lie algebra is n.

3. **Groups with Abelian and degenerate Lie rings.** It is obvious from the definition of the p-Lie ring L of a group G that if G is Abelian, then L is also Abelian. The converse is not true, however, even if we restrict ourselves to p-groups. The group of order 16, type V, (i) in the tables of Burnside[3] has an Abelian Lie algebra which is p-isomorphic to the Lie algebra of the group of order 16 which is Abelian and of type III, (ii) in the tables mentioned above. It is seen, therefore, as indicated in the introduction, that non-isomorphic p-groups may have p-isomorphic Lie algebras. It should be mentioned here, however, that all Abelian p-groups of the same order have the same Lie algebra in the sense of ordinary isomorphism.

Before reducing our problem to the study of p-groups, we must point out that there exist groups whose p-Lie ring is identically zero, or "degenerate". This occurs when and only when $G_m = G$ for all m. The p-Lie algebra of a prime-power group is non-degenerate if and only if p divides the order of the group. Hence a p-group G has one and only one non-degenerate p-Lie algebra, and so we are justified in calling this simply the Lie algebra of G.

It will be seen that there exist finite groups which have a degenerate p-Lie algebra for some p, yet which possess subgroups whose p-Lie algebras are not degenerate. Hence, for arbitrary groups, the study of the p-Lie algebras of subgroups is not only difficult, but probably fruitless.

4. **Reduction to p-groups.** It is obvious that the p-Lie ring of a finite group has a finite basis over $GF(p)$, and so is a Lie algebra. There are also infinite groups for which this is true. In this section we shall show that if L is the p-Lie *algebra* of any group, it is also the p-Lie algebra of a p-group.

The following theorems follow directly from the definition of the p-Lie algebra of a group and from the isomorphism theorems for groups. A proof of the first theorem is given for the sake of rigor.

THEOREM 1. *If H is an invariant subgroup of G contained in all the groups G_n, then the p-Lie ring of G/H is p-isomorphic to that of G.*

Proof. Let $\bar{G}$ denote the group G/H, and $\bar{Z}_i$ the corresponding subgroups of $\bar{G}$. It is known that $\bar{Z}_i \cong Z_i/Z_i \cap H$, where the symbol $\cong$ denotes isomorphism. Similarly,

$$\bar{G}_n \cong G_n/G_n \cap H \cong G_n \cdot H/H.$$

When H is contained in all of the G_n, we have

$$\bar{G}_{(i)} \cong \bar{G}_i/\bar{G}_{i+1} \cong (G_i \cdot H/H)/(G_{i+1} \cdot H/H) \cong G_i/G_{i+1} \cong G_{(i)},$$

so the p-Lie rings of G and $\bar{G}$ are direct sums of the same linear spaces. It remains to be shown that the Lie multiplication and p-operation are the same in each.

[3] W. Burnside, *Theory of Groups of Finite Order*, 2d ed., Cambridge, 1911, pp. 145–146.

If $g \epsilon G$, we let $\bar{g}$ represent the coset of g mod H. Then δ_n, defined by

$$\delta_n(g_n \cdot G_{n+1}) = \bar{g}_n \cdot \bar{G}_{n+1},$$

is an isomorphic mapping of G_n/G_{n+1} onto $\bar{G}_n/\bar{G}_{n+1}$.

We can now define an isomorphism $\bar{d}_n$ between $\bar{G}_n/\bar{G}_{n+1}$ and $\bar{G}_{(n)}$ by

$$\bar{d}_n = d_n \delta_n^{-1},$$

where d_n is the isomorphism between G_n/G_{n+1} and $G_{(n)}$.

Then if

$$d_n(g_n \cdot G_{n+1}) = y_n \epsilon G_{(n)} = \bar{G}_{(n)},$$

we have

$$\bar{d}_n(\bar{g}_n \cdot \bar{G}_{n+1}) = y_n \epsilon \bar{G}_{(n)}.$$

Hence

$$\begin{aligned} \bar{d}_{i+j}([\bar{g}_i, \bar{g}_j] \cdot \bar{G}_{i+j+1}) &= d_{i+j}\delta_{i+j}^{-1}(\overline{[g_i, g_j]} \cdot \bar{G}_{i+j+1}) \\ &= d_{i+j}([g_i, g_j] \cdot G_{i+j+1}). \end{aligned}$$

The two Lie rings are therefore isomorphic in the ordinary sense.

Finally,

$$\begin{aligned} \bar{d}_{np}(\bar{g}_n^p \cdot \bar{G}_{np+1}) &= d_{np}\delta_{np}^{-1}(\overline{g_n^p} \cdot \bar{G}_{np+1}) \\ &= d_{np}(g_n^p \cdot G_{np+1}), \end{aligned}$$

so the two Lie rings are p-isomorphic.

THEOREM 2. *Given any group G, the p-Lie ring of G/G_n for any G_n is p-isomorphic to the quotient ring L/L_n, where L is the p-Lie ring of G and*

$$L_n = G_{(n)} + G_{(n+1)} + \cdots.$$

The proof of this theorem goes through exactly as does the proof of Theorem 1.

If an arbitrary group G has for a given p a p-Lie ring with a finite basis over $GF(p)$, it is obvious that the groups G_n must all be equal after a certain one, say G_m. Since $GF(p)$ has only a finite number of elements, the Lie algebra L must also have only a finite number of elements, so G/G_m must be finite. From Theorem 1, since G_m is contained in all the G_n, the group G/G_m must have a p-Lie algebra which is p-isomorphic to L. This is the first step in our reduction, namely, that if a Lie *algebra* arises from any group, it also arises from a finite group.

Given a finite group G, let us denote by p^α the highest power of p which divides the order of any element of G.

LEMMA. *The set G^{p^α} consists of all elements of order prime to p, and so contains all of the q-Sylow groups of G for q prime to p.*

Proof. Suppose that g is an element of order n, prime to p. There exist integers A, B such that

$$An + Bp^{\alpha} = 1.$$

Then $g = g^{An+Bp^{\alpha}} = g^{Bp^{\alpha}} = (g^{B})^{p^{\alpha}}$, so $G^{p^{\alpha}}$ contains all elements of order prime to p.

Conversely, suppose that g is an element of order $m = rp^{k}$, where r is equal to 1 or prime to p, and

$$1 \leqq k \leqq \alpha.$$

Then $(g^{r})^{p^{k}} = 1$, and so $(g^{r})^{p^{\alpha}} = 1$. Hence $(g^{p^{\alpha}})^{r} = 1$, and so $g^{p^{\alpha}}$ is of order r or some factor of r, say 1; at any rate, its order is prime to p. Hence $G^{p^{\alpha}}$ contains only elements of order prime to p.

It follows from the proof of this lemma that $G^{p^{\alpha}}$ raised to any power of p must be equal to itself, so $G^{p^{\alpha}}$ and hence also $\{G^{p^{\alpha}}\}$ must be contained in all of the groups G_n.

In a finite group G, the groups Z_i must all be equal after a certain one, which we shall designate by Z_{μ}. The group Z_{μ} is contained in all the G_n, and for nilpotent groups $Z_{\mu} = 1$, by definition.

THEOREM 3. *If G is a finite group, the p-Lie algebra of G is p-isomorphic to the p-Lie algebra of the p-Sylow group of the nilpotent group G/Z_{μ}.*

Proof. If G is nilpotent, there is only one Sylow group for each prime factor of the order of G, each of these is invariant, and G is equal to their direct product. Let $p, p_1, p_2, \cdots, p_n$ be the prime factors of the order of G, and $S(p)$, $S(p_1)$, $S(p_2), \cdots, S(p_n)$ the corresponding Sylow groups. Let S denote the direct product $S(p_1) \times S(p_2) \times \cdots \times S(p_n)$. We have shown that $\{G^{p^{\alpha}}\}$ contains all of the $S(p_i)$, and so $S \subseteq \{G^{p^{\alpha}}\}$. Hence the p-Lie algebra of G/S is p-isomorphic to that of G. But $G = S(p) \times S$, so we have

$$G/S = S(p).$$

Hence by Theorem 1, the p-Lie algebra of G is the same as that of $S(p)$, the p-Sylow group of G.

If G is not nilpotent, we know that G/Z_{μ} is nilpotent and has the same p-Lie algebra as does G.

This theorem completes the proof that a Lie algebra arises from a p-group if it arises from any group whatsoever. It also gives a criterion for the non-degeneracy of the p-Lie algebra of a finite group, namely, that p divide the order of G/Z_{μ}. Finally, it proves the statement made before that a finite group G (say one whose order is divisible by p) may have a degenerate p-Lie algebra (when G/Z_{μ} has order prime to p) and yet have a subgroup (its p-Sylow group, for instance) whose p-Lie algebra is not degenerate.

5. **The counter example.** We are now in a position to prove the statement (B) of the introduction.

In order that a Lie algebra L should be the p-Lie algebra of a group, it has been seen necessary that L be a restricted nilpotent Lie algebra over $GF(p)$, whose p-operation satisfies conditions (1), (2). It is obvious that we must also have the condition $L^{p^j} = 0$ for some j.

The following Lie algebra will show that these conditions are not sufficient. Let L be of order 4 over $GF(2)$ (with basis elements x_1, x_2, x_3, x_4), defined by

$$[x_1, x_2] = x_3, \qquad x_1^2 = x_2^2 = x_3, \qquad x_3^2 = x_4,$$

with all other commutators and p-th powers of basis elements equal to zero. When the p-operation satisfies the conditions (1), (2), it is shown by Zassenhaus that it is uniquely determined by its effect on a basis, so we have defined L completely.

If we let the notation (x, y) stand for the linear space spanned by the elements x and y, we have

$$L_1 = (x_1, x_2, x_3, x_4), \qquad L_2 = (x_3, x_4), \qquad L_3 = (x_4).$$

Then

$$\begin{aligned} L &= L_{(1)} + L_{(2)} + L_{(3)} \\ &= (x_1, x_2) + (x_3) + (x_4). \end{aligned}$$

In this Lie algebra all higher commutators are zero, and it can be shown that

$$(x + y)^p = x^p + y^p + [x, y],$$

for p equal to 2. Hence in L,

$$(x_1 + x_2)^2 = x_3 + x_3 + x_3 = x_3.$$

Since $x_3^2 = x_4$, every element of $L_{(1)}$ must have a non-zero eighth power, so if L is the Lie algebra of a group G, there must be in G an element of order at least 8. We have proved that if G exists we can assume it to be a group of order 16. Of the groups of order 16, there are four with elements of order 8. These are given here with their generators and defining relations, numbered as in the tables of Burnside:

(i) $P^8 = 1, \quad Q^2 = 1, \quad Q^{-1}PQ = P^5,$

(vii) $P^8 = 1, \quad Q^2 = 1, \quad Q^{-1}PQ = P^7,$

(viii) $P^8 = 1, \quad Q^2 = 1, \quad Q^{-1}PQ = P^3,$

(ix) $P^8 = 1, \quad Q^4 = 1, \quad Q^{-1}PQ = P^7, \quad Q^2 = P^4.$

Some calculation will show that for each of these groups, G_4 consists of the elements 1, P^4, and $G_5 = 1$.

From Theorem 2, the group $\bar{G}$ defined by $\bar{G} = G/G_4$ must have the Lie algebra $\bar{L}$ defined by $\bar{L} = L/L_3$, which has basis elements y_1, y_2, y_3 and is defined by

$$[y_1, y_2] = y_3, \qquad y_1^2 = y_2^2 = y_3,$$

with all other commutators and p-th powers of basis elements equal to zero.

If L is to be the Lie algebra of any of the four groups of order 16 given above, $\bar{L}$ must be the Lie algebra of a group of order 8 isomorphic to G/G_4, where G is one of the groups of order 16. There are two groups of order 8, and it is readily seen that $\bar{L}$ is the Lie algebra of only one of them, the group whose generators and defining relations are given by

$$P^4 = 1, \qquad Q^4 = 1, \qquad Q^{-1}PQ = P^{-1}, \qquad Q^2 = P^2.$$

It is obvious that this group is not isomorphic to G/G_4 ($G_4 = 1, P^4$) for any of the groups G of order 16. Hence the Lie algebra L which we have constructed is not the p-Lie algebra of any group.

PRINCETON UNIVERSITY.

CONVERGENCE AND DIVERGENCE OF NON-HARMONIC GAP SERIES

By M. Kac

1. Introduction. The theory of convergence of trigonometric (harmonic) series with gaps of Hadamard's type is by now a closed chapter. However, new difficulties present themselves when one tries to get a parallel theory for series

$$\sum_{k=1}^{\infty} a_k e^{i\lambda_k t} \quad \lambda_k > 0; \lambda_{k+1}/\lambda_k > q > 1 \qquad (k = 1, 2, \cdots), \tag{1}$$

without assuming that the λ's are integers. The source of these difficulties is the fact that the non-harmonic series must necessarily be considered on the infinite interval $-\infty < t < \infty$. In the harmonic case the terms have the same period and the problem is simplified both by having to deal with a finite interval and by the properties of orthogonality. To meet the new situation methods of proof had to be modified and some new devices invented. It should, however, be mentioned that the solution obtained is not complete. In §3 we prove that if $\sum |a_k|^2 = \infty$ and $\lambda_{k+1}/\lambda_k > q > (5^{\frac{1}{2}} + 1)/2$ the series (1) diverges almost everywhere, whereas in the harmonic case it suffices to assume that $\lambda_{k+1}/\lambda_k > q > 1$. Thus far we have not been able to remove the condition involving $(5^{\frac{1}{2}} + 1)/2$.

2. Convergence of non-harmonic gap series.

Theorem 1. *If $\sum |a_k|^2$ converges, the series* (1) *converges almost everywhere for $-\infty < t < \infty$.*

The method of proof is a modification of a method of Marcinkiewicz.[1] Let $0 < \delta < \lambda_2 - \lambda_1$. Since $\lambda_{k+1} - \lambda_k \geqq \lambda_2 - \lambda_1$ $(k \geqq 1)$, the intervals $(\lambda_k - \delta, \lambda_k)$ do not overlap and we can define a function $f(x)$ by putting

$$f(x) = a_k \qquad \text{for } \lambda_k - \delta < x \leqq \lambda_k,$$
$$f(x) = 0 \qquad \text{otherwise.}$$

Obviously $f(x) \epsilon L^2(-\infty, \infty)$ and hence its Fourier transform is $(C, 1)$ summable almost everywhere.[2] In particular, it follows that

$$\lim_{n \to \infty} \int_0^{\lambda_n} \left(1 - \frac{x}{\lambda_n}\right) f(x) e^{itx}\, dx \tag{2}$$

Received May 25, 1941.

[1] J. Marcinkiewicz, *A new proof of a theorem on Fourier series*, Journal of the London Mathematical Society, vol. 8(1933), p. 279.

[2] See, for instance, E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, Oxford, 1937, pp. 84–85. The theorem in question is due to Plancherel.

exists almost everywhere. I say now that for every t

$$\lim_{n\to\infty}\int_0^{\lambda_n}\frac{x}{\lambda_n}f(x)e^{itx}\,dx = 0. \tag{3}$$

In fact

$$\left|\int_0^{\lambda_n}\frac{x}{\lambda_n}f(x)e^{itx}\,dx\right| = \left|\sum_{k=1}^{n}\frac{a_k}{\lambda_n}\int_{\lambda_k-\delta}^{\lambda_k}xe^{itx}\,dx\right| \leqq \sum_{k=1}^{n}\frac{|a_k|}{\lambda_n}\int_{\lambda_k-\delta}^{\lambda_k}x\,dx$$

$$\leqq \delta\sum_{k=1}^{n}\frac{|a_k|\lambda_k}{\lambda_n} \leqq \delta\sum_{k=1}^{n}\frac{|a_k|}{q^{n-k}};$$

and since $a_k \to 0$ as $n \to \infty$, it follows immediately that

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{|a_k|}{q^{n-k}} = 0.$$

(2) and (3) imply that

$$\lim_{n\to\infty}\int_0^{\lambda_n}f(x)e^{itx}\,dx$$

exists almost everywhere. But

$$\int_0^{\lambda_n}f(x)e^{itx}\,dx = \frac{1-e^{-i\delta t}}{it}\sum_{k=1}^{n}a_k e^{i\lambda_k t},$$

and since $(1 - e^{-i\delta t})/it$ differs from 0 except on a denumerable set, we get that

$$\lim_{n\to\infty}\sum_{k=1}^{n}a_k e^{i\lambda_k t}$$

exists for almost every t. This completes the proof of Theorem 1.

3. Divergence of non-harmonic gap series.

THEOREM 2. *If $\{\lambda_k\}$ is a sequence of positive numbers satisfying the gap condition*

$$\frac{\lambda_{n+1}}{\lambda_n} \geqq q > \frac{1+5^{\frac{1}{2}}}{2}$$

and $\sum |a_k|^2 = \infty$, then the series (1) *diverges almost everywhere on* $(-\infty, \infty)$.

For the special case in which the λ's are integers and the partial sums accordingly have period 2π, the result was obtained by A. Zygmund under less restrictive conditions.[3] We shall adopt his method with a modification which seems essential.

[3] A. Zygmund, *Trigonometrical Series*, Monografje Matematyczne, Warsaw, 1935, pp. 120–122.

The method of Zygmund leads to the following general theorem which will be stated as a lemma.

LEMMA 1. *Let Ω be a set with a completely additive measure $\mu(E)$ which has the property that the measure of the whole set Ω is 1 and the measure of each set consisting of a single element is 0. Let $\{\varphi_j(p)\}$ be a sequence of functions defined on Ω such that*

(i) *a constant K exists such that for every measurable set $E \subset \Omega$*

$$\int_E^* |\varphi_j(p)|^2 \, d\mu \geqq K \cdot \mu(E) \qquad (j = 1, 2, \cdots)$$

(the star superscript indicates that the integral is taken with respect to the measure μ in Ω),

(ii) *$\varphi_j(p)$ is an orthonormal set on Ω, and, moreover,*

(iii) *$\varphi_j(p)\varphi_k(p)$ $(j, k = 1, 2, \cdots; j \neq k)$ is an orthonormal set on Ω.*

Then if $\sum |a_k|^2 = \infty$ the series $\sum a_k\varphi_k(p)$ diverges almost everywhere on Ω.

We omit the proof since it can be copied step by step from that of Zygmund (loc. cit.).

The main idea of our proof will consist in introducing a completely additive measure on $(-\infty, \infty)$ so as to make Lemma 1 applicable. Let D denote the class of absolutely continuous distribution functions $\sigma(t)$, with absolutely continuous inverses, for which $\sigma(-\infty) = 0$ and

$$\int_{-\infty}^{+\infty} d\sigma(t) = 1. \tag{4}$$

LEMMA 2. *If $\{\mu_k\}$ is a sequence of real numbers for which $\sum \mu_k^{-2} < \infty$, then there exists a distribution function $\sigma(t) \in D$ such that the μ's are zeros of the Fourier-Stieltjes transform of $\sigma(t)$.*

It follows from a result of Wintner[4] that if $\sum b_k^2 < \infty$ the distribution function $\sigma(t)$ defined by the equation

$$e^{-u^2} \prod_{k=1}^{\infty} \cos b_k u = \int_{-\infty}^{+\infty} e^{iut} \, d\sigma(t) \tag{5}$$

is an analytic function and therefore $\sigma(t)$ is absolutely continuous and has an absolutely continuous inverse. Putting $u = 0$ we see that (4) holds. Thus $\sigma(t) \in D$.

Now, put $b_k = \pi/2\mu_k$ and observe that the μ's are zeros of the left member of (5). This proves Lemma 2.

LEMMA 3. *If $\lambda_{n+1}/\lambda_n \geqq q > (5^{\frac{1}{2}} + 1)/2$ then*

$$\sum_{m,n,j,k=1}^{\infty}{}^{*} \frac{1}{(\lambda_m + \lambda_n - \lambda_j - \lambda_k)^2} < \infty,$$

[4] *On analytic convolutions of Bernoulli distributions*, American Journal of Mathematics, vol. 56(1934), pp. 659–663.

where the star on the summation sign indicates that the terms for which m, n, j, k are not distinct should be rejected.

The following proof of an even more general statement is due to R. P. Agnew. Let $p > 0$ and let S, finite or infinite, be defined by

$$S = \sum_{m,n,j,k=1}^{\infty}{}^{*} \frac{1}{|\lambda_m + \lambda_n - \lambda_j - \lambda_k|^p}.$$

To estimate S, we arrange the terms of the series in such a way as to combine the terms for which the greatest of the indices m, n, j, k has a fixed value N. It thus appears that

$$S \leqq 4 \sum_{N=4}^{\infty} \sum_{1 \leqq \alpha,\beta,\gamma < N} \frac{1}{|\lambda_N + \lambda_\alpha - \lambda_\beta - \lambda_\gamma|^p}.$$

From the gap condition it follows immediately that $\lambda_N - \lambda_\beta - \lambda_\gamma > 0$ and therefore

$$S \leqq 4 \sum_{N=4}^{\infty} \sum_{1 \leqq \alpha,\beta,\gamma < N} \frac{1}{(\lambda_N - \lambda_{N-1} - \lambda_{N-2})^p}$$
$$\leqq 4 \sum_{N=4}^{\infty} \frac{(N-1)^3}{(\lambda_N - \lambda_{N-1} - \lambda_{N-2})^p}.$$

Using the gap condition $\lambda_{n+1}/\lambda_n \geqq q > (5^{\frac{1}{2}} + 1)/2$ again and letting Q denote the positive number defined by

$$Q = 1 - \frac{1}{q} - \frac{1}{q^2},$$

we see that

$$\lambda_N - \lambda_{N-1} - \lambda_{N-2} > Q\lambda_N > Q\lambda_1 q^{N-1}$$

and consequently

$$S \leqq \frac{4}{(Q\lambda_1)^p} \sum_{N=4}^{\infty} \frac{(N-1)^3}{(q^p)^{N-1}} < \infty.$$

This proves Lemma 3.

Proof of Theorem 2. Let Ω be the infinite set $(-\infty, \infty)$. It is a trivial consequence of the gap condition that

$$\sum_{m \neq j} \frac{1}{(\lambda_m - \lambda_j)^2} < \infty. \tag{6}$$

Let $\mu_1, \mu_2, \mu_3, \cdots$ denote in some order the points of the denumerable set of elements representable in the form $\lambda_m - \lambda_j$ with $m \neq j$ and the points of the denumerable set of elements representable in the form $\lambda_m + \lambda_n - \lambda_j - \lambda_k$ with m, n, j and k distinct. Then by Lemma 3 and (6) $\sum \mu_k^{-2} < \infty$, and the distri-

bution function $\sigma(t)$ can be constructed as in Lemma 2. To a Lebesgue measurable set $E \subset \Omega$ is now assigned a new measure

$$\int_E d\sigma(t),$$

the integral being the ordinary Lebesgue-Stieltjes integral. This measure obviously satisfies the conditions of Lemma 1 and, moreover, sets of measure 0 according to this new measure are precisely the same as sets of ordinary Lebesgue measure 0.

Let a star superscript designate integrals formed with respect to the new measure. Put $\varphi_j(t) = e^{i\lambda_j t}$ and notice that the requisite conditions of orthogonality

$$\int^* \varphi_m(t)\overline{\varphi_j(t)}\,dt = \int_{-\infty}^{+\infty} e^{i\lambda_m t} e^{-i\lambda_j t}\,d\sigma(t) = \int_{-\infty}^{+\infty} e^{i(\lambda_m-\lambda_j)t}\,d\sigma(t) = 0$$

and

$$\int^* \varphi_m(t)\varphi_n(t)\overline{\varphi_j(t)}\,\overline{\varphi_k(t)}\,dt = \int_{-\infty}^{+\infty} e^{i(\lambda_m+\lambda_n-\lambda_j-\lambda_k)t}\,d\sigma(t) = 0$$

are satisfied since the numbers of the form $\lambda_m - \lambda_j$ and $\lambda_m + \lambda_n - \lambda_j - \lambda_k$ are zeros of the Fourier-Stieltjes transform of $\sigma(t)$. The condition (i) of Lemma 1 is trivially satisfied with $K = 1$ and the normality is obvious in view of (4). Thus, Lemma 1 is applicable and the proof of Theorem 2 is completed.

4. Another application of the method.

THEOREM 3. *If*

$$\sum_{j\neq k} \frac{1}{(\lambda_j - \lambda_k)^2} < \infty \tag{7}$$

and, for some $\epsilon > 0$, $\sum |a_k|^{2-\epsilon} < \infty$, *the series* $\sum a_k e^{i\lambda_k t}$ *converges almost everywhere.*

Let the μ's be numbers of the form $\lambda_j - \lambda_k$ $(j \neq k)$ and construct $\sigma(t)$ as in Lemma 2. Let $\varphi_k(x) = e^{i\lambda_k \sigma^{-1}(x)}$. Then

$$\int_0^1 \varphi_j(x)\overline{\varphi_k(x)}\,dx = \int_{-\infty}^{+\infty} e^{i(\lambda_j-\lambda_k)t}\,d\sigma(t) = 0$$

and

$$\int_0^1 |\varphi_k(x)|^2\,dx = \int_{-\infty}^{+\infty} d\sigma(t) = 1.$$

The φ_k's form an orthonormal system and by a theorem of Menchoff[5] $\sum a_k\varphi_k(x)$ converges almost everywhere. This in view of the fact that $\sigma^{-1}(x)$ is absolutely continuous implies our theorem.

CORNELL UNIVERSITY.

[5] D. Menchoff, *Sur les séries de fonctions orthogonales*, Fundamenta Mathematicae, vol. 11(1928), pp. 375–420, Theorem 12.

LAPLACE TRANSFORMS OF MULTIPLY MONOTONIC FUNCTIONS

By Norman N. Royall, Jr.

1. **Introduction.** Let $s = \sigma + i\tau$ be a complex variable and let t be a real variable. Then if the integral in the right member of the equation

$$f(s) = \int_0^\infty e^{-st}\alpha(t)\,dt \tag{1.1}$$

converges, $f(s)$ is called the Laplace transform of $\alpha(t)$. The fundamental properties of the transformation (1.1) have been studied by Widder, Doetsch, and others (see [8], [2]). The region of convergence of the integral in the right member of (1.1) is a half-plane $\sigma > \sigma_0$ and $f(s)$ is an analytic function of s there.

In this paper we are concerned with the effect upon the structure of $f(s)$ of the assumption of various degrees of monotonic order on $\alpha(t)$. To this end we have the following

Definition of Monotonic Order. A real function $\beta(t)$ of a real variable t is said to be monotonic of order k in the interval $0 < t < \infty$ if the function satisfies the condition $(-1)^n\beta^{(n)}(t) \geqq 0$ $(n = 0, 1, 2, \cdots, k)$ in that interval.

A similar definition applies for the interval $0 \leqq t < \infty$.

Condition A. From the above definition it is clear that $\alpha(t)$ may be monotonic in character and the integral in (1.1) fail to converge as is shown by the function $\alpha(t) = 1/t^2$. Hence, throughout this paper we shall assume that, in addition to the specified monotonic character, $\alpha(t)$ and its derivatives, when these latter are assumed to exist, are such that

$$\int_0^\infty e^{-st}\alpha^{(n)}(t)\,dt$$

converges for $\sigma > 0$. Functions $\alpha(t)$ which satisfy this requirement will be said to "satisfy condition A".

Definition of $f_R(s)$. In the statement of the theorems it is convenient to use the symbol $f_R(s)$ which is defined as the integral

$$f_R(s) = \int_0^R e^{-st}\alpha(t)\,dt.$$

2. **Transforms of triply monotonic functions.**

Theorem 1. *If $\alpha(t)$ is monotonic of order three on the interval $0 < t < \infty$ and satisfies condition* A, *then*

Received April 15, 1941. A part of a thesis presented at Brown University in partial fulfillment of the requirements for the degree of Doctor of Philosophy; prepared under the direction of Professor J. D. Tamarkin.

$$\left|\frac{f(s) - f_{R_1}(s)}{e^{-sR_1}}\right| \geqq \left|\frac{f(s) - f_{R_2}(s)}{e^{-sR_2}}\right| \tag{2.1}$$

for $\sigma > 0$ *and* $R_1 < R_2$.

Proof. We have

$$f(s) - f_R(s) = \int_R^\infty e^{-st}\alpha(t)\,dt$$

and by the substitution $T = t - R$ we obtain

$$\frac{f(s) - f_R(s)}{e^{-sR}} = \int_0^\infty e^{-sT}\alpha(T + R)\,dT$$

and thus

$$\frac{f(s) - f_R(s)}{e^{-sR}} = \int_0^\infty e^{-\sigma t}\alpha(t + R)\,d\int_0^t e^{-i\tau T}\,dT. \tag{2.2}$$

Integrating the right member of (2.2) twice by parts using the Stieltjes integral formula for convenience gives

$$\frac{f(s) - f_R(s)}{e^{-sR}} = \int_0^\infty S_1(\tau, t)K(\sigma, t, R)\,dt, \tag{2.3}$$

where

$$S_1(\tau, t) = \int_0^t dt_1 \int_0^{t_1} e^{-i\tau T}\,dT \tag{2.4}$$

and

$$K(\sigma, t, R) = e^{-\sigma t}[\alpha''(t + R) - 2\sigma\alpha'(t + R) + \sigma^2\alpha(t + R)]. \tag{2.5}$$

Separating $S_1(\tau, t)$ into its real and imaginary parts and denoting these by $S_1^r(\tau, t)$ and $S_1^i(\tau, t)$ respectively, we obtain

$$S_1^r(\tau, t) = \frac{1 - \cos \tau t}{\tau^2} \tag{2.6}$$

and

$$S_1^i(\tau, t) = \frac{\sin \tau t - \tau t}{\tau^2}. \tag{2.7}$$

Now clearly the right member of (2.6) is non-negative for all $t \geqq 0$ and, since $|\sin x| \leqq |x|$, it follows that for a fixed τ the right member of (2.7) is of one sign.

But $K(\sigma, t, R) \geqq 0$; indeed the hypothesis of triple monotony insures that each of its three terms is non-negative. Moreover, for a fixed value of σ the expression $K(\sigma, t, R)$ decreases steadily as R increases. Hence, separating the

right member of (2.3) into its real and imaginary parts, we obtain

$$\frac{f(s) - f_R(s)}{e^{-sR}} = \int_0^\infty S_1^r(\tau, t)K(\sigma, t, R)\,dt + i\int_0^\infty S_1^i(\tau, t)K(\sigma, t, R)\,dt \tag{2.8}$$

and we observe that the absolute values of both the real and the imaginary parts of the right member of (2.8) diminish steadily as R increases. The theorem follows from this fact.

3. **Transforms of doubly monotonic functions.** If the hypothesis on $\alpha(t)$ is reduced to the assumption that it is monotonic of order two, it is possible to obtain a similar, but somewhat weaker, result than in the previous theorem. The result is given precisely in

THEOREM 2. *If $\alpha(t)$ is monotonic of order two on the interval* $0 < t < \infty$ *and satisfies condition* A, *then*

$$|f(s) - f_{R_1}(s)| \geqq |f(s) - f_{R_2}(s)| \tag{3.1}$$

for $\sigma > 0$ and $R_1 < R_2$.

Proof. As in the previous theorem, integrating twice by parts we obtain (2.8). But from the present hypothesis we have only $\alpha(t) \geqq 0$, $\alpha'(t) \leqq 0$, and $\alpha''(t) \geqq 0$; hence although we can guarantee the positive character of $K(\sigma, t, R)$, we cannot assure here its monotone character as R increases, for now nothing is known concerning the monotone character of $\alpha''(t + R)$ as R increases. For this reason we now focus our attention upon the *first* factor in each integrand in the right member of (2.8).

By the change of variable $t' = t + R$ we obtain from (2.8) after suppressing the accent

$$|f(s) - f_R(s)| = \left| \int_R^\infty S_1^r(\tau, t - R)K(\sigma, t)\,dt + i\int_R^\infty S_1^i(\tau, t - R)K(\sigma, t)\,dt \right|, \tag{3.2}$$

where now

$$K(\sigma, t) = e^{-\sigma t}[\alpha''(t) - 2\sigma\alpha'(t) + \sigma^2\alpha(t)]. \tag{3.3}$$

We next consider the *square* of the members of (3.2); for it is clear that if $|f(s) - f_R(s)|^2$ is a monotone non-increasing function of R, then the members of (3.2) are also. But this square is

$$\int_R^\infty S_1^r(\tau, t - R)K(\sigma, t)\,dt \int_R^\infty S_1^r(\tau, t' - R)K(\sigma, t')\,dt'$$
$$+ \int_R^\infty S_1^i(\tau, t - R)K(\sigma, t)\,dt \int_R^\infty S_1^i(\tau, t' - R)K(\sigma, t')\,dt'.$$

In the above expression both $K(\sigma, t)$ and $K(\sigma, t')$ are independent of R; hence to prove that this expression is monotonic it suffices to prove that

$$S_1^r(\tau, t - R)S_1^r(\tau, t' - R) + S_1^i(\tau, t - R)S_1^i(\tau, t' - R) \tag{3.4}$$

is monotonic.

It is important to observe here that both terms of (3.4) are non-negative for any fixed τ and that t and t' may have any two fixed values whatsoever, subject merely to $t \geqq R$ and $t' \geqq R$. In particular, we do not necessarily assume that $t = t'$.

Now let

$$\tau(t - R) = x; \qquad \tau(t' - R) = y. \tag{3.5}$$

Thus we are led to consider the expression

$$H(R) = (1 - \cos x)(1 - \cos y) + (\sin x - x)(\sin y - y). \tag{3.6}$$

The first term of the right member of (3.6) is non-negative; moreover, we observe from (3.5) that x and y have the same sign, hence the two factors of the second term are alike in sign. Hence $H(R) \geqq 0$. Differentiating with respect to R gives

$$H'(R) = -\tau[(1 - \cos x)y + (1 - \cos y)x]$$

since from (3.5) we have

$$\frac{dx}{dR} = -\tau = \frac{dy}{dR}.$$

But since $t \geqq R$ and $t' \geqq R$ we have $\operatorname{sgn} x = \operatorname{sgn} \tau = \operatorname{sgn} y$. Hence the sign of the bracket above is the same as the sign of τ; therefore if $\tau > 0$, the bracket is positive and $H'(R) < 0$; if $\tau < 0$, the bracket is negative, and again $H'(R) < 0$. Hence $H(R)$ is non-increasing; but this implies that the expression (3.4) is non-increasing and the theorem follows.

It might be supposed that (3.1) would be valid if $\alpha(t)$ were merely a simple monotonic non-increasing function. This is, however, *not* the case. The author has given an example[1] of such a function for which the inequality (3.1) fails to hold.

As an immediate consequence of Theorem 2 we have

THEOREM 3. *If $\alpha(t)$ is monotonic of order two on the interval $0 < t < \infty$ and satisfies condition* A, *then*

$$| f_R(s) | \leqq 2 | f(s) | \tag{3.7}$$

for $\sigma > 0$ and $R \geqq 0$.

Proof. From Theorem 2 by letting $R_1 \to 0$ and dropping the subscript from R_2 we obtain

$$| f(s) - f_R(s) | \leqq | f(s) |.$$

Hence

$$| f_R(s) | - | f(s) | \leqq | f(s) |,$$

and (3.7) follows from this fact.

[1] Thesis, Brown University, May 1940.

From Theorem 1 we can obtain in like manner for the half-plane $\sigma \geqq \sigma_0 > 0$ the somewhat better result

$$|f_R(s)| \leqq (1 + e^{-\sigma R}) \, |f(s)|.$$

If $\alpha(t)$ is completely monotonic on the interval $0 < t < \infty$, the inequality (2.1) can be obtained readily by use of the Bernstein-Widder theorem (see footnote 1).

An Application to Cornu's Spiral.[2] The equations expressing the coördinates of points on this spiral parametrically as functions of R are

$$x = \frac{a}{2^{1/2}} \int_0^R \frac{\sin t}{t^{1/2}} \, dt; \qquad y = \frac{a}{2^{1/2}} \int_0^R \frac{\cos t}{t^{1/2}} \, dt.$$

The origin of coördinates is given by $R = 0$ and by allowing R to become infinite one obtains the asymptotic points of the spiral, namely,

$$\left(\pm \frac{a\pi^{1/2}}{2^{1/2}}, \pm \frac{a\pi^{1/2}}{2^{1/2}} \right).$$

Now consider

$$f(s) = \frac{a}{2^{1/2}} \int_0^\infty e^{-st} t^{-1/2} \, dt.$$

We have to do here with $\alpha(t) = a2^{-1/2}t^{-1/2}$ which is completely monotonic on $0 < t < \infty$; i.e., which has derivatives of all orders there satisfying the requirement $(-1)^n \alpha^{(n)}(t) \geqq 0$.

Taking $\sigma = 0$ and $\tau = 1$, we find

$$f - f_R = \frac{a}{2^{1/2}} \int_R^\infty e^{-it} t^{-1/2} \, dt$$

$$= \frac{a}{2^{1/2}} \int_R^\infty \frac{\cos t}{t^{1/2}} \, dt - i \frac{a}{2^{1/2}} \int_R^\infty \frac{\sin t}{t^{1/2}} \, dt,$$

and the absolute values of both the real and imaginary parts here are monotone non-increasing functions of R. But the right member of the above expression may be arranged so as to display the asymptotic points thus:

$$\left[\frac{a}{2^{1/2}} \int_0^\infty \frac{\cos t}{t^{1/2}} \, dt - \frac{a}{2^{1/2}} \int_0^R \frac{\cos t}{t^{1/2}} \, dt \right] - i \left[\frac{a}{2^{1/2}} \int_0^\infty \frac{\sin t}{t^{1/2}} \, dt - \frac{a}{2^{1/2}} \int_0^R \frac{\sin t}{t^{1/2}} \, dt \right].$$

Thus we see that the distance from a point on the spiral to an asymptotic point decreases steadily as the point moves along the spiral. It is of course not necessary that a curve have this property in order to have the spiral property.

[2] For a drawing of this beautiful curve see the *Encyclopaedia Britannica*, 14th edition, vol. 6, article on "Special Curves", section 37 and sketch on page 893.

4. Univalent transforms of triply monotonic functions.

THEOREM 4. *If $\alpha(t)$ is monotonic of order three on the interval $0 \leqq t < \infty$ and is not identically zero, then*

$$f(s) = \int_0^\infty e^{-st}\alpha(t)\,dt$$

is univalent in the half-plane of convergence $\sigma > 0$, and vanishes only at infinity.

Proof. Let $w_0 = f(s_0)$, where $s_0 = \sigma_0 + i\tau_0$ is any finite point in the half-plane $\sigma > 0$. Let C be a contour enclosing s_0, rectangular and symmetrical to the σ-axis, and with corners at the points $c - id$, $d - id$, $d + id$, and $c + id$, where c is any number such that $0 < c < \sigma_0$ and d is taken large enough for the contour C to enclose s_0.

Now consider the number of zeros of $f(s) - w_0$ inside C. To prove univalence it is sufficient to show that the number of zeros of $f(s) - w_0$ inside C is *exactly unity* no matter how large d is taken; and since the number of zeros of a function in a domain is at least as great as in any subdomain of the former, it suffices to consider the number of zeros as $d \to \infty$.

First observe that if n denotes the number of zeros of $f(s) - w_0$ in $\sigma > c$, then n is certainly a *positive integer*, for there is *at least one zero there.* We now show that there is *exactly* one.

Let N be the number of zeros of $f(s) - w_0$ inside C, then a well-known formula of function theory is

$$N = \frac{1}{2\pi i}\int_C \frac{f'(s)}{f(s) - w_0}\,ds.$$

This may be expressed as the sum of four integrals along the sides of the rectangle which we denote as α, β, γ, and δ; referring to the sides $(c - id, d - id)$, $(d - id, d + id)$, $(d + id, c + id)$, and $(c + id, c - id)$ respectively. Thus

$$N = \frac{1}{2\pi i}\int_\alpha \frac{f'(s)}{f(s) - w_0}\,ds + \frac{1}{2\pi i}\int_\beta \frac{f'(s)}{f(s) - w_0}\,ds + \frac{1}{2\pi i}\int_\gamma \frac{f'(s)}{f(s) - w_0}\,ds + \frac{1}{2\pi i}\int_\delta \frac{f'(s)}{f(s) - w_0}\,ds.$$

Consider the behavior of the first three integrals as $d \to \infty$. Integrating twice by parts, we have

$$f(s) = \int_0^\infty \frac{1 - \cos \tau t}{\tau^2} K(\sigma, t)\,dt + i\int_0^\infty \frac{\sin \tau t - \tau t}{\tau^2} K(\sigma, t)\,dt, \tag{4.1}$$

where $K(\sigma, t)$ is defined as in (3.3). From this we see that the integrand in the first integral of the right member of (4.1) is positive (even if $\tau = 0$) since $\alpha(t) \not\equiv 0$. Therefore $f(s_0) = w_0 \neq 0$. We now show that

$$\left|\frac{f'(s)}{f(s) - w_0}\right| = O(d^{-2})$$

on α, β, and γ as $d \to \infty$. We have

$$f(s) = \int_0^\infty e^{-\sigma t} e^{-i\tau t} \alpha(t)\, dt,$$

$$|f(s)| \leqq \alpha(0+) \int_0^\infty e^{-\sigma t}\, dt$$

$$= \frac{\alpha(0+)}{\sigma} \to 0$$

as $\sigma \to \infty$, and hence $f(s) \to 0$ on β as $d \to \infty$.

But on α and γ we have

$$f(s) = e^{-\sigma t} \alpha(t) \left. \frac{e^{-i\tau t}}{-i\tau} \right]_0^\infty + \frac{1}{i\tau} \int_0^\infty e^{-i\tau t}[-\sigma e^{-\sigma t} \alpha(t) + e^{-\sigma t} \alpha'(t)]\, dt,$$

$$|f(s)| \leqq \left| \frac{\alpha(0+)}{\tau} \right| + \left| \frac{\alpha(0+)}{\tau} \right| \int_0^\infty \sigma e^{-\sigma t}\, dt + \left| \frac{\alpha'(0+)}{\tau} \right| \int_0^\infty e^{-\sigma t}\, dt,$$

$$|f(s)| \leqq 2 \frac{\alpha(0+)}{d} + \frac{1}{c} \frac{|\alpha'(0+)|}{d},$$

and hence $f(s) \to 0$ on α and γ also as $d \to \infty$.

We next examine $f'(s)$ on α, β, and γ as $d \to \infty$. Denoting the real and imaginary parts of the right member of (4.1) by $U(\sigma, \tau)$ and $V(\sigma, \tau)$ respectively, we have

$$f'(s) = \frac{\partial V}{\partial \tau} - i \frac{\partial U}{\partial \tau}.$$

But

$$\frac{\partial}{\partial \tau} \left(\frac{1 - \cos \tau t}{\tau^2} \right) = \frac{\tau t \sin \tau t - 2 + 2 \cos \tau t}{\tau^3}$$

and

$$\frac{\partial}{\partial \tau} \left(\frac{\sin \tau t - \tau t}{\tau^2} \right) = \frac{\tau t \cos \tau t + \tau t - 2 \sin \tau t}{\tau^3}.$$

Therefore,

$$f'(s) = \frac{1}{\tau^3} \int_0^\infty (\tau t \cos \tau t + \tau t - 2 \sin \tau t) K(\sigma, t)\, dt$$
$$- \frac{i}{\tau^3} \int_0^\infty (\tau t \sin \tau t - 2 + 2 \cos \tau t) K(\sigma, t)\, dt.$$

But since $K > 0$, we have

$$|f'(s)| \leqq \frac{3}{\tau^2} \int_0^\infty t K(\sigma, t)\, dt + \frac{6}{|\tau^3|} \int_0^\infty K(\sigma, t)\, dt.$$

Now let

$$M = \max [\alpha''(0+), -\alpha'(0+), \alpha(0+)];$$

then

$$K(\sigma, t) \leqq Me^{-\sigma t}(1 + 2\sigma + \sigma^2),$$

and we obtain

$$\int_0^\infty tK(\sigma, t)\, dt \leqq M\left(\frac{1}{\sigma^2} + \frac{2}{\sigma} + 1\right).$$

Similarly

$$\int_0^\infty K(\sigma, t)\, dt \leqq M\left(\frac{1}{\sigma} + 2 + \sigma\right).$$

Hence on α and γ

$$|f'(s)| \leqq \frac{3}{d^2} M\left(\frac{1}{c^2} + \frac{2}{c} + 1\right) + \frac{6}{d^3} M\left(\frac{1}{c} + 2 + d\right).$$

Thus finally

$$|f'(s)| = O(d^{-2})$$

as $d \to \infty$.

But on β we have

$$f(s) = \int_0^\infty e^{-st}\,\alpha(t)\, dt,$$

$$f'(s) = \int_0^\infty -te^{-st}\alpha(t)\, dt,$$

$$|f'(s)| \leqq M \int_0^\infty te^{-st}\, dt,$$

$$|f'(s)| \leqq M\left(\frac{1}{\sigma^2}\right),$$

$$|f'(s)| = O(d^{-2})$$

as $d \to \infty$. Hence the fraction

$$\left|\frac{f'(s)}{f(s) - w_0}\right| = O(d^{-2})$$

on α, β, and γ. Therefore,

$$\left|\int_\alpha \frac{f'(s)}{f(s) - w_0}\, ds\right| = [d - c]O(d^{-2}) \to 0,$$

$$\left|\int_\beta \frac{f'(s)}{f(s) - w_0}\, ds\right| = [2d]O(d^{-2}) \to 0,$$

$$\left|\int_\gamma \frac{f'(s)}{f(s) - w_0}\, ds\right| = [d - c]O(d^{-2}) \to 0.$$

It remains only to discuss the integral along δ as $d \to \infty$. We now have

$$n = \lim_{d \to \infty} \frac{1}{2\pi i} \int_\delta \frac{f'(s)}{f(s) - w_0}\, ds$$

$$= \frac{1}{2\pi i} \int_{c+i\infty}^{c-i\infty} \frac{f'(s)}{f(s) - w_0}\, ds$$

$$= \frac{1}{2\pi i} \log \{f(s) - w_0\} \Big]_{c+i\infty}^{c-i\infty}.$$

But $\log \{f(s) - w_0\} = \log | f(s) - w_0 | + i \operatorname{am} \{f(s) - w_0\}$ and since $f(c + i\infty) = f(c - i\infty) = 0$ we have

$$n = \frac{1}{2\pi} \Delta_C \operatorname{am} \{f(s) - w_0\},$$

where Δ_C is the variation along the line $\sigma = c$ from $c + i\infty$ to $c - i\infty$. But clearly this is equivalent to

$$n = \frac{1}{2\pi} \Delta_{C'} \operatorname{am} \{w - w_0\},$$

where C' is the image of the line $\sigma = c$ under the transformation $w = f(s)$. We now examine the map of this line $\sigma = c$.

From (4.1) we observe that for $\sigma > 0$ we have $U > 0$ and $\operatorname{sgn} V = -\operatorname{sgn} \tau$; thus the half-plane $\sigma > 0$ maps onto the half-plane $U > 0$, the first and fourth quadrants being interchanged and the axis $\tau = 0$ mapping onto the axis $V = 0$. Moreover, since $U(\sigma, \tau)$ does *not* change sign with τ, whereas $V(\sigma, \tau)$ *does* change sign with τ, we see that the line $\sigma = c$ maps into a curve which is symmetrical with respect to the U-axis.

But from

$$U = \int_0^\infty e^{-\sigma t} \alpha(t) \cos \tau t\, dt$$

we find that

$$-\frac{dU}{d\tau} = \int_0^\infty e^{-\sigma t} \alpha(t)\, t \sin \tau t\, dt,$$

and integrating *three* times by parts gives

$$\frac{dU}{d\tau} = \int_0^\infty M_2(\tau, t) e^{-\sigma t} [\alpha'''(t) - 3\sigma\alpha''(t) + 3\sigma^2\alpha'(t) - \sigma^3\alpha(t)]\, dt, \tag{4.2}$$

where

$$\begin{aligned} M_2(\tau, t) &= \int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} T \sin \tau T\, dT \\ &= \frac{1}{2!} \int_0^t (t - T)^2 T \sin \tau T\, dT. \end{aligned} \tag{4.3}$$

Evaluating the right member of (4.3) by elementary methods gives

$$2!\,M_2 = \frac{2t \cos \tau t}{\tau^3} + \frac{4t}{\tau^3} - \frac{6 \sin \tau t}{\tau^4}. \tag{4.4}$$

A glance at (4.3) shows that $M_2 = 0$ when $\tau = 0$. But we shall show that $M_2 > 0$ when $\tau > 0$; for we have

$$\tau^4 2!M_2 = 2\tau t \cos \tau t + 4\tau t - 6 \sin \tau t,$$

and letting $\tau t = x$, whence $x > 0$ when $\tau > 0$, we consider the expression

$$G(x) = 2x \cos x + 4x - 6 \sin x, \qquad G(0) = 0. \tag{4.5}$$

Using the crudest estimate, we see from the fact that $-2x + 4x - 6 > 0$ according as $2x > 6$ that $G(x) > 0$ if $x > 3$. But by successive differentiation we have

$$\begin{aligned} G'(x) &= 4 - 2x \sin x - 4 \cos x, & G'(0) &= 0,\\ G''(x) &= 2 \sin x - 2x \cos x, & G''(0) &= 0,\\ G'''(x) &= 2x \sin x, & G'''(0) &= 0. \end{aligned}$$

Now clearly $G'''(x) > 0$ when $0 < x < \pi$; hence $G''(x)$ is increasing, and thus positive, on that interval; hence $G'(x)$ is increasing and therefore positive there; so finally $G(x) > 0$ for $0 < x < \pi$. Combining this result with the one above, we see that $G(x) > 0$ for *all* $x > 0$.

Therefore $M_2 > 0$ when $\tau > 0$, and so in the integrand of the right member of (4.2) the first factor is *positive* and from the hypothesis on $\alpha(t)$ the second factor is *negative*; hence $dU/d\tau < 0$ when $\tau > 0$. But since $U > 0$ when $\tau > 0$, this implies that U is steadily diminishing as τ increases. We see now that the image of the line $\sigma = c$ is therefore a *simple closed curve without double points.* Since this curve has been shown to be symmetrical with respect to the U-axis, it is not necessary to consider $\tau < 0$.

We are now in a position to see that the number of zeros of $f(s) - w_0$ in the half-plane $\sigma > c$ is *exactly* unity, for n was given by

$$n = \frac{1}{2\pi} \Delta_{C'} \operatorname{am} \{w - w_0\}$$

and if w_0 is *outside* C' and w traverses C', then $n = 0$; on the other hand if w_0 is *inside* C', $n = -1$ if w traverses C' negatively and $n = +1$ if w traverses C' positively. But, as previously remarked, n is certainly a *positive* integer and hence is $+1$, the curve being traversed in the positive manner as τ varies from $+\infty$ to $-\infty$.

For the foregoing arguments it is essential that s_0 be a finite point in the half-plane $\sigma > 0$, but it has been shown that $f(s)$ does not vanish at any such point, and since $f(\infty) = 0$ as is easily seen from the foregoing arguments, it follows that $f(s)$ is univalent in $\sigma > 0$ and vanishes only at infinity.

5. A counter example for transforms of doubly monotonic functions. The conclusion of univalence does *not* remain valid when the hypothesis on $\alpha(t)$ is reduced from monotone of order three to monotone of order two. For consider the function

$$\alpha(t) = e^{-t}\left(\tfrac{5}{2} + \sin t\right). \tag{5.1}$$

Here $\alpha(t) > 0$, and by successive differentiation we obtain

$$\alpha'(t) = -e^{-t}\left(\tfrac{5}{2} + \sin t - \cos t\right),$$

$$\alpha''(t) = e^{-t}\left(\tfrac{5}{2} - 2\cos t\right),$$

$$\alpha'''(t) = -e^{-t}\left(\tfrac{5}{2} - 2\sin t - 2\cos t\right),$$

hence $\alpha'(t) < 0$ and $\alpha''(t) > 0$. But since we have $\max | -2\sin t - 2\cos t | = 2\cdot 2^{\frac{1}{2}} > \frac{5}{2}$ it is clear that $\alpha'''(t)$ is of variable sign. Therefore, $\alpha(t)$ is monotone of order two on the interval $0 \leqq t < \infty$ but not of order three.

With this choice of $\alpha(t)$ we have

$$f(s) = \int_0^\infty e^{-st} e^{-t}\left(\tfrac{5}{2} + \sin t\right) dt,$$

$$= \frac{5}{2}\int_0^\infty e^{-(s+1)t}\, dt + \int_0^\infty e^{-(s+1)t} \sin t\, dt \tag{5.2}$$

$$= \frac{5}{2}\cdot\frac{1}{s+1} + \frac{1}{1+(s+1)^2}.$$

Hence

$$f'(s) = -\frac{\frac{5}{2}[1+(s+1)^2]^2 + 2(s+1)^3}{(s+1)^2[1+(s+1)^2]^2}. \tag{5.3}$$

We now show that the numerator of the right member of (5.3) has a *zero with positive real part* for which the denominator does not also vanish. From this fact and a well-known argument based on Rouche's theorem[3] it will follow that $f(s)$ cannot be univalent in $\sigma > 0$. Expanding this numerator, we are led to examine the zeros of

$$5s^4 + 24s^3 + 52s^2 + 52s + 24. \tag{5.4}$$

It is easy to verify that (5.4) has no *real* zeros. Placing $s = iy$ in (5.4) we obtain

$$5y^4 - 52y^2 + 24 + i(-24y^3 + 52y), \tag{5.5}$$

an expression whose real and imaginary parts never vanish simultaneously; hence (5.4) has no *purely imaginary* zeros.

[3] See, for example, E. C. Titchmarsh, *The Theory of Functions*, Oxford, 1932, p. 198.

We apply the formula

$$2\pi N = \Delta_C \operatorname{am} \{f(z)\} \tag{5.6}$$

to the polynomial (5.4) with the contour C consisting of a part of the real axis from $s = 0$ to $s = R$, thence around a quadrant of the circle with center at origin and radius R to the positive imaginary axis, and finally down the latter axis to the origin. The variation of the amplitude of (5.4) along the portion of C consisting of the real axis is, of course, zero. On the quadrant of the circle we have for sufficiently large R,

$$\Delta \operatorname{am} \{5s^4 + 24s^3 + \cdots + 24\} = \Delta \operatorname{am} \{5R^4e^{4i\theta}\} + \Delta \operatorname{am} \{1 + O(R^{-1})\} \tag{5.7}$$

and since θ varies from zero to $\frac{1}{2}\pi$ on this arc, the variation of the amplitude there is $2\pi + O(R^{-1})$.

Finally, the variation along the imaginary axis is given, as we see from (5.5), by

$$\operatorname{am} (5s^4 + 24s^3 + \cdots + 24) = \operatorname{arc} \tan \left[\frac{-24y^3 + 52y}{5y^4 - 52y^2 + 24}\right]. \tag{5.8}$$

The roots of the numerator in the bracket are $-\left(\frac{13}{6}\right)^{1/2}$, 0, and $+\left(\frac{13}{6}\right)^{1/2}$, whereas the positive roots of the denominator are $+8^{1/2}$ and $+\left(\frac{12}{5}\right)^{1/2}$. Hence as y varies from $+\infty$ to zero we have the following scheme to display the variation of the inverse tangent

$$\begin{array}{llllllllll} y\colon & +\infty, & & +8^{1/2}, & & +\left(\frac{12}{5}\right)^{1/2}, & & \left(\frac{13}{6}\right)^{1/2}, & & 0. \\ \operatorname{arc}\tan\colon & 0, & -, & -\infty, & +, & +\infty, & -, & 0, & +, & 0. \end{array}$$

Hence the amplitude oscillates as follows: beginning at zero, moving negatively to $-\frac{1}{2}\pi$, then into the third quadrant but *reversing and returning* to $-\frac{1}{2}\pi$, back to zero, on into the first quadrant but *reversing again*, returning finally to zero. Hence the variation of the amplitude along the imaginary axis is zero.

Thus the only contribution made to the variation is made by the arc, and since this latter is 2π, we see that $f'(s)$ has a zero inside the contour C if R is sufficiently large.

The argument here depends essentially upon the non-interlacing of the zeros of the numerator and denominator of the rational fraction in (5.8). This depends upon the fact that $\frac{12}{5} > \frac{13}{6}$, and this in turn upon the original choice of the constant $\frac{5}{2}$ in the counter example. If this original constant is taken large enough to make $\alpha(t)$ monotonic of order three, the zeros interlace, the variation along the imaginary axis is -2π, hence on the entire contour is zero, and the argument above fails.

Bibliography

1. S. Bernstein, *Sur les fonctions absolument monotones*, Acta Mathematica, vol. 52(1929), pp. 1–66.
2. G. Doetsch, *Theorie und Anwendung der Laplace Transformation*, Berlin, 1937.
3. L. Fejér, *Potenzreihe mit mehrfach monotoner Koeffizientenfolge und ihre Legendrepolynome*, Proceedings of the Cambridge Philosophical Society, vol. 31(1935), pp. 307–316.
4. L. Fejér, *Trigonometrische Reihen und Potenzreihen mit mehrfach monotoner Koeffizientenfolge*, Transactions of the American Mathematical Society, vol. 39(1936), pp. 18–59.
5. L. Fejér and G. Szegö, *Über die monotone Konvergenz von Potenzreihen mit mehrfach monotoner Koeffizientenfolge*, Prace Matematyczno-Fizyczne, vol. 44(1937), pp. 15–25.
6. P. Montel, *Leçons sur les fonctions univalentes*, Paris, 1933.
7. G. Szegö, *Some recent investigations concerning the sections of power series and related developments*, Bulletin of the American Mathematical Society, vol. 42(1936), pp. 505–522.
8. D. V. Widder, *A generalization of Dirichlet's series and of Laplace's integrals by means of a Stieltjes integral*, Transactions of the American Mathematical Society, vol. 31(1929), pp. 694–743.
9. D. V. Widder, *Necessary and sufficient conditions for the representation of a function as a Laplace integral*, Transactions of the American Mathematical Society, vol. 33(1931), pp. 851–892.

Winthrop College.

POWER SERIES WITH MULTIPLY MONOTONIC SEQUENCES OF COEFFICIENTS

By G. Szegö

Introduction

1. In various papers, L. Fejér[1] dealt with the following interesting theorem:

A. *Let the sequence* $\{a_n\}$ *be monotonic of order* 4. *Then the power series* $f(z) = \sum_{n=0}^{\infty} a_n z^n$ *is regular and univalent for* $|z| < 1$.

Here and in what follows, a sequence $\{a_n\}$ is called monotonic of order k if all the differences

$$\Delta^{(\nu)} a_n = a_n - \binom{\nu}{1} a_{n+1} + \binom{\nu}{2} a_{n+2} - \cdots + (-1)^{\nu} \binom{\nu}{\nu} a_{n+\nu} \tag{1}$$

are non-negative for $\nu = 0, 1, 2, \cdots, k$; $n = 0, 1, 2, \cdots$. If we write

$$s_n^{(3)}(z) = \binom{n+3}{3} + \binom{n+2}{3} z + \binom{n+1}{3} z^2 + \cdots + z^n, \tag{2}$$

the representation

$$f(z) = \sum_{n=0}^{\infty} \Delta^{(4)} a_n \cdot s_n^{(3)}(z) + (1 - z)^{-1} \cdot \lim_{n \to \infty} a_n \tag{3}$$

holds provided the latter limit exists. This is, for instance, the case if $\Delta^{(0)} a_n \geqq 0$, $\Delta^{(1)} a_n \geqq 0$.

Obviously, the expression

$$\sum_{n=0}^{\infty} A_n s_n^{(3)}(z) + A(1 - z)^{-1}, \qquad A_n \geqq 0, A \geqq 0, \tag{4}$$

furnishes a parametric representation of the class of power series mentioned above.

The proof of Fejér is based on the fact that $\Re s_n^{(3)}(z)$ is decreasing when $z = re^{i\theta}$, $0 < r < 1$, and θ increases from 0 to π. This property can be extended without difficulty to non-negative linear combinations of the $s_n^{(3)}(z)$, that is,

Received May 28, 1941; presented to the American Mathematical Society, September 12, 1940.

[1] (a) *Trigonometrische Reihen und Potenzreihen mit mehrfach monotoner Koeffizientenfolge*, Transactions of the American Mathematical Society, vol. 39(1936), pp. 18–59. (b) *Hatványsorok többszörösen monoton együtthatósorozattal*, Matematikai és Természettudományi Értesitö, vol. 55(1936), pp. 1–29. (c) *Untersuchungen über Potenzreihen mit mehrfach monotoner Koeffizientenfolge*, Acta Litterarum ac Scientiarum, vol. 8(1936), pp. 89–115.

because of the representation (3) to the most general power series of the considered kind.

2. The simple example $f(z) = 1 + z + z^2 + \cdots + z^{n-1}$ given by Fejér shows that the order 4 required in the theorem above cannot be reduced to 1. The intermediate orders 2 and 3, however, seem to be more difficult. In the present paper I prove:

B. *Theorem* A *remains true for monotonic sequences of order* 3.

C. *Theorem* A *is not true for monotonic sequences of order* 2.

Regarding these questions I exchanged several letters with Professor Fejér during previous years. I communicated to him the result C (with a proper counter example) in a letter of August 7, 1936. In his answer (September 18, 1936) he informed me that E. Egerváry proved in the case of Theorem B that $f'(z) \neq 0$ in the unit circle $|z| = 1$. Also, at a later time, S. Sidon[2] gave a proof of C different from mine.

3. Of course, the representation (3) can be extended to an arbitrary order $\nu \geqq 1$ [see Fejér, loc. cit., (c), p. 104, formula (4)]

$$(5)\qquad \begin{cases} f(z) = \sum_{n=0}^{\infty} \Delta^{(\nu)} a_n \cdot s_n^{(\nu-1)}(z) + (1-z)^{-1} \cdot \lim_{n\to\infty} a_n, \\ s_n^{(\nu-1)}(z) = \binom{n+\nu-1}{\nu-1} + \binom{n+\nu-2}{\nu-1} z + \cdots + z^n, \end{cases}$$

so that a parametric representation similar to that mentioned above holds true for the class of power series with a sequence of coefficients monotonic of a given order ν. The proof of B is based on certain properties of $s_n^{(2)}(z)$ which can be again extended to non-negative linear combinations.

The expressions $s_n^{(\nu-1)}(z)$ are connected in a well-known fashion with the Cesàro sums of the geometric series. We have

$$(6)\qquad s_n^{(\nu)}(z) = s_0^{(\nu-1)}(z) + s_1^{(\nu-1)}(z) + \cdots + s_n^{(\nu-1)}(z).$$

Proof of Theorem B

1. Let $n \geqq 1$, and

$$(7)\qquad s_n^{(2)}(e^{i\varphi}) = x_n(\varphi) + i y_n(\varphi).$$

If we denote by φ_0 the angle determined by the conditions $\sin^2 \frac{1}{2}\varphi_0 = 0.7$, $\frac{1}{2}\pi < \varphi_0 < \pi$, the following inequalities hold:

$$(8)\qquad y_n(\varphi) > 0, \qquad 0 < \varphi < \pi,$$

$$(9)\qquad x_n'(\varphi) < 0, \qquad 0 < \varphi \leqq \varphi_0,$$

$$(10)\qquad y_n'(\varphi) < 0, \qquad \tfrac{1}{2}\pi < \varphi \leqq \pi,$$

$$(11)\qquad x_n(\varphi_1) - x_n(\varphi_2) > 0, \qquad 0 \leqq \varphi_1 \leqq \tfrac{1}{2}\pi;\ \varphi_0 \leqq \varphi_2 \leqq \pi.$$

[2] S. Sidon, *Über Potenzreihen mit monotoner Koeffizientenfolge*, Acta Litterarum ac Scientiarum, vol. 9(1940), pp. 244–246.

Obviously, $\varphi_0 > \frac{1}{2}\pi$. Also, considering the case $n = 1$, we notice that the lower bound $\frac{1}{2}\pi$ in (10) cannot be replaced by a smaller one. In the present section we give the very simple proofs of these inequalities.

Before doing this, we observe that inequalities (8)–(11) lead to corresponding inequalities for the real and imaginary parts of a function of the form

$$f_n(z) = \sum_{m=0}^{n} A_m s_m^{(2)}(z) + A(1 - z)^{-1}$$

on the unit circle $z = e^{i\varphi}$, provided $A_m \geqq 0$, $A \geqq 0$, and $f_n(z)$ is not a constant. This is clear for the first term of $f_n(z)$; for the second term we have to take into account that $(1 - e^{i\varphi})^{-1} = \frac{1}{2} + \frac{1}{2}i \cot \frac{1}{2}\varphi$. [If $A_\nu = 0$ $(\nu = 1, 2, \cdots, n; A > 0)$, the sign $<$ in (9) and the sign $>$ in (11) have to be replaced by $=$.]

But this result implies the univalency of the mapping $w = f_n(z)$ for $|z| \leqq 1$. Indeed, because of (8) we can confine ourselves to the values $0 < \varphi < \pi$. Let $0 < \varphi_1 < \varphi_2 < \pi$. Then, $f_n(e^{i\varphi_1}) - f_n(e^{i\varphi_2})$ cannot vanish if both φ_1 and φ_2 belong to the range in (9), or to the range in (10). Similarly, it cannot vanish if φ_1 and φ_2 satisfy the conditions laid down in (11).

Now let $A_m = \Delta^{(4)} a_m$, $A = \lim_{n\to\infty} a_n$. Since $f(z) = \lim_{n\to\infty} f_n(z)$ uniformly for $|z| \leqq r$, $r < 1$, the univalency of $f(z)$ also follows. This will prove Theorem B.

2. We start with the important inequality $\mathfrak{I} s_n^{(1)}(e^{i\varphi}) > 0$ due to Lukács;[3] this implies [cf. (6)] (8).

3. Before we proceed to the proof of the inequalities (9)–(11), the following formulas might be mentioned:

$$x_n(\varphi) = \tfrac{1}{4}(n + 1)(n + 2) + \tfrac{1}{8}(\sin \tfrac{1}{2}\varphi)^{-3}\{(2n + 3) \sin \tfrac{1}{2}\varphi - \sin (n + \tfrac{3}{2})\varphi\}, \tag{12}$$

$$\begin{aligned} x_n'(\varphi) &= \tfrac{1}{8} \cos \tfrac{1}{2}\varphi\, (\sin \tfrac{1}{2}\varphi)^{-3} \\ &\quad \cdot \left\{-2n - 3 - (n + \tfrac{3}{2}) \frac{\cos (n + \frac{3}{2})\varphi}{\cos \frac{1}{2}\varphi} + \frac{3}{2} \frac{\sin (n + \frac{3}{2})\varphi}{\sin \frac{1}{2}\varphi}\right\}, \end{aligned} \tag{13}$$

$$\begin{aligned} y_n'(\varphi) &= \tfrac{1}{8}(\sin \tfrac{1}{2}\varphi)^{-2} \\ &\quad \cdot \left\{-(n^2 + 3n + 3) + 3\left(\frac{\sin (n + 1)\frac{1}{2}\varphi}{\sin \frac{1}{2}\varphi}\right)^2 - n \frac{\sin (n + \frac{3}{2})\varphi}{\sin \frac{1}{2}\varphi}\right\}. \end{aligned} \tag{14}$$

They can easily be verified by means of the representation

$$\begin{aligned} s_n^{(2)}(z) &= \tfrac{1}{2} \sum_{m=0}^{n} (n + 2 - m)(n + 1 - m)z^m \\ &= \tfrac{1}{2}(n + 1)(n + 2)(1 - z)^{-1} - z(1 - z)^{-3}\{n + 1 - (n + 2)z + z^{n+2}\}. \end{aligned} \tag{15}$$

[3] Cf. L. Fejér, *Einige Sätze, die sich auf das Vorzeichen einer ganzen rationalen Funktion beziehen*, Monatshefte für Math. und Physik, vol. 35(1928), pp. 305–344; especially pp. 336–337.

4. Since $x_n(\varphi)$ is a cosine polynomial with positive coefficients, inequality (9) is trivial for $\varphi < \pi/n$. Suppose $\pi/(n + \frac{3}{2}) \leqq \varphi \leqq \varphi_0$; for these values of φ we write the statement in the form

$$\frac{3}{2n+3} \frac{\sin (n + \frac{3}{2})\varphi}{\sin \frac{1}{2}\varphi} - \frac{\cos (n + \frac{3}{2})\varphi}{\cos \frac{1}{2}\varphi} < 2 \tag{16}$$

which follows from

$$\left(\frac{3}{2n+3}\right)^2 \frac{1}{\sin^2 \frac{1}{2}\varphi} + \frac{1}{\cos^2 \frac{1}{2}\varphi} < 4. \tag{17}$$

The left side is a function of $\sin^2 \frac{1}{2}\varphi$ which is convex from downward, so that it suffices to prove (17) for the end-values of φ. The corresponding inequalities are

$$\left(\frac{3}{2n+3}\right)^2 \left(\sin \frac{\pi}{2n+3}\right)^{-2} + \left(\cos \frac{\pi}{2n+3}\right)^{-2} < 4, \qquad \left(\frac{3}{2n+3}\right)^2 \frac{10}{7} + \frac{10}{3} < 4.$$

Both left-hand side expressions are decreasing when n increases. Writing $n = \frac{1}{2}$ in the first one, we obtain $\frac{25}{8}$, writing $n = 1$ in the second one, we obtain $\frac{404}{105}$. This proves (9).

5. Inequality (10) is a consequence of

$$3 (\sin \tfrac{1}{2}\varphi)^{-2} + n (\sin \tfrac{1}{2}\varphi)^{-1} < n^2 + 3n + 3. \tag{18}$$

This is true for $\varphi = \frac{1}{2}\pi$, $n = 2$, therefore also for $\frac{1}{2}\pi \leqq \varphi \leqq \pi$, $n \geqq 2$, since $(\sin \frac{1}{2}\varphi)^{-1} < 3$. But $y_1(\varphi) = \sin \varphi$, so that the statement holds for all $n \geqq 1$.

6. Inequality (11) is trivial for $n = 1$. In case $n = 2$ we obtain

$$x_2(\varphi) = 6 + 3 \cos \varphi + \cos 2\varphi.$$

This is decreasing for $0 \leqq \varphi < \varphi'$ and increasing for $\varphi' < \varphi \leqq \pi$, where $\cos \varphi' = -\frac{3}{4}$. Now $\cos \varphi_0 = -0.4$ so that $\frac{1}{2}\pi < \varphi_0 < \varphi'$. But $x_2(\frac{1}{2}\pi) = 5 > \max \{x_2(\varphi_0), x_2(\pi)\} = \max \{4.12, 4\}$, and (11) is proved in this case.

Let $n \geqq 3$. Since $x_n(\varphi)$ is decreasing for $0 \leqq \varphi \leqq \pi/n$, we can assume that $\varphi_1 > \pi/n$. On account of (12) it suffices to show that

$$(2n + 3)(\sin \tfrac{1}{2}\varphi_1)^{-2} - (\sin \tfrac{1}{2}\varphi_1)^{-3} > (2n + 3)(\sin \tfrac{1}{2}\varphi_2)^{-2} + (\sin \tfrac{1}{2}\varphi_2)^{-3}. \tag{19}$$

The function of $\sin \frac{1}{2}\varphi_1$ on the left side is decreasing for $\sin \frac{1}{2}\varphi_1 > \frac{3}{2}(2n + 3)^{-1}$; this condition is satisfied when $\varphi_1 > \pi/n$. Therefore its minimum is attained for $\varphi_1 = \frac{1}{2}\pi$. The function on the right side attains its maximum for $\varphi_2 = \varphi_0$. Thus, it remains to show that

$$2(2n + 3) - 2^{\frac{3}{2}} > \tfrac{10}{7}(2n + 3) + (\tfrac{10}{7})^{\frac{3}{2}}.$$

This is indeed true for $n \geqq 3$.

This makes the proof of Theorem B complete.

Proof of Theorem C

1. The counter example which proves C will have the form

$$f(z) = ps_m^{(1)}(z) + qs_m^{(1)}(z), \tag{20}$$

where $p > 0$, $q > 0$, m and n integers, and p, q, m, n have to be chosen properly; we prove that for a proper z, $| z | < 1$, the derivative $f'(z)$ vanishes.

Let λ be fixed; obviously,

$$\begin{aligned} \lim_{n\to\infty} n^{-2} s_n^{(1)}(e^{-\lambda/n}) &= \lim_{n\to\infty} n^{-2} \sum_{m=0}^{n} (n + 1 - m)e^{-m\lambda/n} = \int_0^1 (1 - t)e^{-\lambda t}\, dt \\ &= \lambda^{-2}(e^{-\lambda} - 1 + \lambda) = S(\lambda); \end{aligned} \tag{21}$$

also

$$\lim_{n\to\infty} n^{-3} s_n^{(1)\prime}(e^{-\lambda/n}) = -S'(\lambda) = \lambda^{-3}\{(2 + \lambda)e^{-\lambda} - 2 + \lambda\}. \tag{22}$$

We shall consider the function $S(\lambda)$ for $\Re\lambda \geqq 0$.

Incidentally, the mapping $\lambda' = S(\lambda)$ is star-shaped for $\Re\lambda > 0$ with respect to $S(0) = \frac{1}{2}$. This follows from the corresponding star-shaped character of $s_n^{(1)}(z)$ in $| z | < 1$ with respect to $s_n^{(1)}(1)$ due to E. Egerváry.[4] The latter statement regarding $s_n^{(1)}(z)$ is of course much deeper than the statement regarding $S(\lambda)$ which can readily be proved directly. However, in what follows, this property is not used.

2. First, we prove that

$$\operatorname{sgn} S'(\tfrac{1}{2}i\pi) = -\operatorname{sgn} S'(\tfrac{5}{2}i\pi) \neq 0. \tag{23}$$

This follows from (22) by simple substitution:

$$\begin{cases} S'(\frac{1}{2}i\pi) = -(\frac{1}{2}\pi)^{-3}(2 - \frac{1}{2}\pi)(1 - i), \\ S'(\frac{5}{2}i\pi) = (\frac{5}{2}\pi)^{-3}(\frac{5}{2}\pi - 2)(1 - i). \end{cases} \tag{24}$$

Thus, if we write $\gamma_0 = 125(4 - \pi)/(5\pi - 4)$,

$$S'(\tfrac{1}{2}i\pi) + \gamma_0 S'(\tfrac{5}{2}i\pi) = 0 \tag{25}$$

holds. Now let

$$\Im\, \frac{S'(\rho e^{i\varphi})}{S'(\frac{1}{2}\pi e^{i\varphi})} = u(\rho, \varphi). \tag{26}$$

Since $u(\rho_0, \varphi_0) = 0$ if $\rho_0 = \frac{5}{2}\pi$, $\varphi_0 = \frac{1}{2}\pi$, we have $u(\rho, \varphi) = 0$ for sufficiently small values of $| \varphi - \varphi_0 |$ and proper $\rho = \rho(\varphi)$, provided

$$\frac{\partial u}{\partial \rho} \neq 0 \qquad \text{for } \rho = \rho_0,\ \varphi = \varphi_0.$$

[4] E. Egerváry, *Abbildungseigenschaften der arithmetischen Mittel der geometrischen Reihe*, Mathematische Zeitschrift, vol. 42(1937), pp. 221–230.

But for $\rho = \rho_0$, $\varphi = \varphi_0$

$$\frac{\partial u}{\partial \rho} = \Im\left\{e^{i\varphi}\frac{S''(\rho e^{i\varphi})}{S'(\frac{1}{2}\pi e^{i\varphi})}\right\} = \Im\,\{i(\tfrac{1}{2}\pi)^3(2 - \tfrac{1}{2}\pi)^{-1}(1 - i)^{-1}S''(\tfrac{1}{2}i\pi)\} \neq 0,$$

since

$$S''(\lambda) = \lambda^{-4}\{(6 + 4\lambda + \lambda^2)e^{-\lambda} - 6 + 2\lambda\},$$
$$S''(\tfrac{1}{2}i\pi) = (\tfrac{1}{2}\pi)^{-4}\{10\pi - 6 + i(-6 + 5\pi + \tfrac{25}{4}\pi^2)\}.$$

3. From the previous result we conclude that the number γ defined by

$$S'(\rho(\varphi)e^{i\varphi}) + \gamma S'(\tfrac{1}{2}\pi e^{i\varphi}) = 0,$$

where $\varphi_0 - \varphi = \frac{1}{2}\pi - \varphi$ is a sufficiently small positive number, is real and positive. We fix φ and write

$$\lambda_1 = \rho(\varphi)e^{i\varphi}, \qquad \lambda_2 = \tfrac{1}{2}\pi e^{i\varphi}, \tag{27}$$

so that

$$S'(\lambda_1) + \gamma S'(\lambda_2) = 0. \tag{28}$$

Here $\Re\lambda_1$, $\Re\lambda_2$, λ_1/λ_2, γ are positive numbers.

4. Now, if we define

$$F_n(z) = s_{[\lambda n]}^{(1)}(z) + \lambda^3\gamma s_n^{(1)}(z), \qquad \lambda = \lambda_1/\lambda_2, \tag{29}$$

we have [cf. (22), (28)]

$$\lim_{n\to\infty} n^{-3}F_n'(e^{-\lambda_2/n}) = -\lambda^3 S'(\lambda_1) - \lambda^3\gamma S'(\lambda_2) = 0. \tag{30}$$

According to Hurwitz's theorem, the function $F_n'(e^{-\lambda/n})$ must vanish if λ is sufficiently near λ_2 and n is sufficiently large. Therefore, $F_n(z)$ cannot be univalent for $|z| < 1$.

The function $F_n(z)$ furnishes the counter example required.

STANFORD UNIVERSITY.

A THEOREM ON DIMENSION

By J. H. Roberts

The present paper grew from an attempt to answer a question of Hurewicz.[1] The answer is given by Theorem 9.1. It is also thought that Theorem 1.1 is of some importance. One application of it, other than Theorem 9.1, is given.

The letter R with positive integral subscript will denote Euclidean space of the indicated dimension.

1. Theorem 1.1. *If the separable metric space M has dimension n, then there is a homeomorphism f of M into a subset of R_{2n+1} such that for any $R_k \subset R_{2n+1}$ $(n + 1 \leqq k \leqq 2n + 1)$ we have* $\dim (f(M)\cdot R_k) \leqq k - n - 1$.

This theorem is shown to be a consequence of the following special case:

Theorem 1.2. *If the separable metric space M has dimension n, then there is a homeomorphism f of M into a subset of R_{2n+1} such that for any $R_{n+1} \subset R_{2n+1}$ we have* $\dim (f(M)\cdot R_{n+1}) \leqq 0$.

We show that the f satisfying the conclusion of Theorem 1.2 will do for Theorem 1.1. The proof is by induction on k. By hypothesis the result is true for $k = n + 1$. Suppose it is true for $k = r < 2n + 1$. Take any R_{r+1} in R_{2n+1}. Select $p \in f(M)\cdot R_{r+1}$. For arbitrary $\epsilon > 0$ there is a rectangular ϵ-domain U of p in R_{r+1} bounded by parts of a finite number of r-dimensional spaces $R_r^1, R_r^2, \cdots, R_r^i$. By our inductive hypothesis, $\dim (f(M)\cdot R_r^j) \leqq r - n - 1$ $(j = 1, 2, \cdots, i)$. Thus $\dim (f(M)\cdot R_{r+1}) \leqq (r - n - 1) + 1 = (r + 1) - n - 1$. This means that our conclusion holds for $k = r + 1$, and the argument is complete.

2. Proof of Theorem 1.2. We first imbed[2] M topologically in a compact space of dimension n. However, to avoid extra terminology we shall continue to use the letter M, which throughout the remainder of this proof will denote a compact n-dimensional metric space.

2.1. Definition of F. Let F denote the space of all continuous mappings of M into subsets of R_{2n+1} with the metric $|f_1 - f_2| = \max_{x \in M} |f_1(x) - f_2(x)|$.

Received June 24, 1941; presented to the American Mathematical Society, April 27, 1940.

[1] W. Hurewicz, *Ueber stetige Bilder von Punktmengen*, Proc. Amsterdam Academy, vol. 30(1927), p. 161, footnote 7. "Es entsteht die Frage ob bei vorgegebenen n^* und n $(n^* > n \geqq 0)$ sich jede n^*-dimensionale separable Menge als ein eindeutiges beiderseits stetiges Bild einer n-dimensionalen Menge mit höchstens $(n^* - n + 1)$-fachen Punkten darstellen lässt."

[2] This is possible, as proved by Hurewicz: *Ueber Einbettung separabler Räume in gleichdimensionale kompakte Räume*, Monatshefte für Math. und Physik, vol. 37(1930), pp. 199–208.

2.2. Definition of F_ϵ. For any $\epsilon > 0$, an element f of F is in F_ϵ provided the following two conditions are satisfied:

2.21. f is an ϵ-mapping; i.e., for $p \in f(M)$ the set $f^{-1}(p)$ has diameter $< \epsilon$;

2.22. for any $R_{n+1} \subset R_{2n+1}$ the set $f(M) \cdot R_{n+1}$ admits a finite, open ϵ-covering $\mathfrak{U}$ of order[3] not exceeding 0; i.e., $\mathfrak{U}$ is a finite collection of non-intersecting sets which are open in R_{2n+1}, each being of diameter less than ϵ.

We make the following assertions:

2.3. F_ϵ is dense in F.

2.4. F_ϵ is open in F.

Then since F is complete, it follows by Baire's theorem[4] that there is a dense G_δ subset of F every element of which is in F_ϵ for every ϵ. Let f be such a mapping. Then from 2.21 it follows that f is a homeomorphism. Take any $R_{n+1} \subset R_{2n+1}$. Then by 2.22 it follows that the set $f(M) \cdot R_{n+1}$ admits a finite open ϵ-covering of order not exceeding 0, for every $\epsilon > 0$. Hence[5] dim $(f(M) \cdot R_{n+1}) \leqq 0$. Thus the proof of Theorem 1.2 will be complete when we have established 2.3 and 2.4.

3. §§3 and 4 are devoted to a proof of 2.3.

Let any $\epsilon > 0$ be given. Let f_1 be any element of F. Let δ_1 and δ_2 be positive numbers. It is convenient to define a function f_2 having certain properties relative to δ_1 and δ_2, and then show that δ_1 and δ_2 can be taken small enough so that the corresponding f_2 is in F_ϵ and is suitably close to f_1. Let $\mathfrak{U} = U_1, U_2, \cdots, U_t$ be an open δ_1-covering of M of order not exceeding n, and for each i select $p_i \in U_i$.

We wish to obtain a particular realization K of the nerve of $\mathfrak{U}$ in R_{2n+1}. We want the complex K to have a special property which does not follow from the ordinary "general position" of its vertices. To indicate what is wanted we consider the case $n = 2$, and suppose K is a geometric 2-complex in R_5, with vertices in "general position". Let R_3 be any 3-space in R_5. Then it follows that at most 4 vertices of K lie in R_3, but there is no bound at all on the number of 1-simplexes of K which *intersect* R_3. We wish to limit the number of intersections of this type. We use the following

Theorem 3.1. *There exists an integer s ($s \leqq (2n + 1)(n + 2)$) such that if $V_1, V_2, \cdots$ is any sequence (finite or countably infinite) of mutually exclusive open sets in R_{2n+1}, then there are points $q_1, q_2, \cdots$, with $q_i \in V_i$, such that if K is any n-dimensional complex on the vertices $q_1, q_2, \cdots$, then for any $R_{n+1} \subset R_{2n+1}$ there are at most s non-intersecting simplexes of K of dimensions not exceeding $n - 1$ which intersect R_{n+1}.*

[3] By the *order* of a finite collection of sets is meant the greatest integer k such that some $k + 1$ of the sets have a non-vacuous intersection.

[4] See Kuratowski's *Topology* I, p. 204.

[5] See Menger, *Dimensionstheorie*, p. 157.

Let $t_1, t_2, \cdots$ be[6] algebraically independent real numbers. For each k let T_k be the set of all numbers of the form $r + t_k$ for r rational. Choose $q_i \in V_i$ so that the j-th coördinate of q_i is in the set T_k, where $k = j + (i - 1)(2n + 1)$. Let K be any n-dimensional complex on $q_1, q_2, \cdots$, and let R_{n+1} be any linear $(n + 1)$-subspace of R_{2n+1}. We write

$$(3.3) \qquad R_{n+1}: x_i = \sum_{j=1}^{n+1} a_{ij}u_j + a_{i0} \qquad (i = 1, 2, \cdots, 2n + 1),$$

where $u_1, u_2, \cdots, u_{n+1}$ are real parameters.

Let m be the smallest integer for which there is a set of real numbers a_1, $a_2, \cdots, a_m$ such that every a_{ij} $(1 \leqq i \leqq 2n + 1, 0 \leqq j \leqq n + 1)$ is algebraically dependent on this set. Clearly $m \leqq (2n + 1)(n + 2)$, and we show that m has the property stated of s in Theorem 3.1, with respect to R_{n+1}.

Suppose that $\sigma_1, \sigma_2, \cdots, \sigma_{m+1}$ are non-intersecting simplexes of K each of dimension less than n, each of which intersects R_{n+1}. For each i $(i = 1, 2, \cdots, m + 1)$ let $b_{i1}, b_{i2}, \cdots, b_{in_i}$ be the set of all coördinates of all the vertices of σ_i. Then the fact that σ_i and R_{n+1} intersect implies that there is for each i at least one value of j, say r, such that b_{ir} is algebraically dependent upon the set consisting of the remaining b_{ij} and the set $a_1, a_2, \cdots, a_m$. Then the set of $m + \sum_{i=1}^{m+1} n_i$ real numbers a_k, b_{ij} $(k = 1, 2, \cdots, m; j = 1, 2, \cdots, n_i; i = 1, 2, \cdots, m + 1)$ is algebraically dependent upon a subset consisting of $m + \sum_{i=1}^{m+1} (n_i - 1) = -1 + \sum_{i=1}^{m+1} n_i$ numbers. Now every b_{ij} differs by a rational number from some t_k, and the totality of such numbers t_k is $\sum_{i=1}^{m+1} n_i$. Then this set of t_k's is algebraically dependent on a set of $-1 + \sum_{i=1}^{m+1} n_i$ real numbers. This is in contradiction to the algebraic independence of $t_1, t_2, \cdots$. Thus Theorem 3.1 is established.

4. We return to the proof of 2.3. Choose $V_i \subset S(f_1(p_i), \delta_2)$ so that $V_1, V_2, \cdots, V_t$ are non-intersecting. Let $q_1, q_2, \cdots, q_t$ be points in R_{2n+1} having the properties stated in Theorem 3.1. Then we have

$$(4.1) \qquad | f_1(p_i) - q_i | < \delta_2 .$$

Let K denote the complex which is the realization of the nerve of $\mathfrak{U}$ on the vertices $q_1, q_2, \cdots, q_t$. Let f_2 be Kuratowski's mapping[7] of M into $\bar{K}$.

We next get a limit on the diameter of the simplexes of K. If q_i and q_j are

[6] For a discussion of algebraic dependence see B. L. van der Waerden's *Moderne Algebra*, especially §64 on *Der Transzendenzgrad*.

[7] See Alexandroff-Hopf, *Topologie*, p. 366. To each point $p \in M$ is assigned a set of numbers $\mu_i(p)$ $(i = 1, 2, \cdots, t)$ by the formula $\mu_i(p) =$ distance from p to $M - U_i$. Then $f_2(p)$ is the center of gravity of a set of masses $\mu_i(p)$ on the points q_i $(i = 1, 2, \cdots, t)$.

in a simplex of K, then $U_i \cdot U_j \neq 0$, and thus $| p_i - p_j | < 2\delta_1$. Then $| q_i - q_j | \leqq | q_i - f_1(p_i) | + | f_1(p_i) - f_1(p_j) | + | f_1(p_j) - q_j |$.

For a limit on the term $| f_1(p_i) - f_1(p_j) |$ we make use of the following function $\eta(\delta)$, connected with the continuous mapping f_1 of the compact space M. For any $\delta > 0$ we write

$$\eta(\delta) = \underset{x,y \in M, |x-y| < \delta}{\text{g.l.b.}} | f_1(x) - f_1(y) |. \tag{4.2}$$

It follows that $\eta(\delta) \to 0$ as $\delta \to 0$.

We can now write

$$| q_i - q_j | \leqq 2\delta_2 + \eta(2\delta_1) = \Delta. \tag{4.3}$$

Since the maximum distance of two points of a simplex is attained for two vertices, it follows that Δ is a bound on the diameter of the simplexes of K.

We next get a bound on $| f_1 - f_2 |$.

For any $p \in M$ we have at least one i such that $p \in U_i$. Then $f_2(p)$ will be in a simplex of K having q_i as a vertex, and $| f_2(p) - q_i | \leqq \Delta$. We have $| f_1(p) - f_2(p) | \leqq | f_1(p) - f_1(p_i) | + | f_1(p_i) - q_i | + | q_i - f_2(p) | \leqq \eta(\delta_1) + \delta_2 + \Delta$. Thus

$$| f_1 - f_2 | \leqq \eta(\delta_1) + \delta_2 + \Delta. \tag{4.4}$$

Now consider the condition that f_2 be an ϵ-mapping. If $f_2(p) = f_2(\bar{p})$, then there is at least one element of $\mathfrak{U}$ which contains both p and $\bar{p}$, whence $| p - \bar{p} | < \delta_1$. Hence f_2 is an ϵ-mapping if $\delta_1 < \epsilon$.

We come finally to the consideration of 2.22 (in the definition of F_ϵ).

Choose any $R_{n+1} \subset R_{2n+1}$. Now $f_2(M) \subset \bar{K}$, hence to show that f_2 has property 2.22 it is sufficient to show that $\bar{K} \cdot R_{n+1}$ admits a finite open ϵ-covering of order not exceeding 0. This follows from the fact, which we shall now prove, that $\bar{K} \cdot R_{n+1}$ admits a finite *closed* ϵ-covering of order not exceeding 0. We show, in fact, that the collection of components of $\bar{K} \cdot R_{n+1}$ is such a closed covering—for suitably chosen δ_1 and δ_2. The finiteness of this collection follows from the finiteness of the complex K.

Let C be a component of $\bar{K} \cdot R_{n+1}$. We obtain a bound on the diameter of C. Suppose $a \in C$, $b \in C$, and $| a - b | > 3(s + 1)\Delta$, where s is the number given in Theorem 3.1 and Δ is defined in 4.3. There is an arc ab in C. Every subarc of ab of diameter not less than Δ must contain at least one point on the boundary of a simplex of K, hence on a simplex of K of dimension not exceeding $n - 1$. Thus there exist points $a_1, a_2, \cdots, a_{s+1}$ on the arc ab such that (1) $(3i - 2)\Delta < | a - a_i | \leqq (3i - 1)\Delta$, and (2) a_i is in a simplex σ_i of K having dimension not exceeding $n - 1$. Now if $i \neq j$, $| a_i - a_j | > 2\Delta$, hence $\sigma_i \cdot \sigma_j = 0$. But then the $s + 1$ mutually exclusive simplexes $\sigma_1, \sigma_2, \cdots, \sigma_{s+1}$, of dimension not exceeding $n - 1$, all intersect C, which is in R_{n+1}, and we have a contradiction to Theorem 3.1. Thus $| a - b | \leqq 3(s + 1)\Delta$; hence $3(s + 1)\Delta$ is an upper bound on the diameter of C.

To complete the proof of 2.3, let $\delta > 0$ be given. We wish to select δ_1 and δ_2 so that $|f_1 - f_2| < \delta$ and $f_2 \in F_\epsilon$. This will be true if the following inequalities hold:

(a) $\eta(\delta_1) + \delta_2 + \Delta = 3\delta_2 + \eta(\delta_1) + \eta(2\delta_1) < \delta$;

(b) $\delta_1 < \epsilon$;

(c) $3(s + 1)\Delta \leqq 3(s + 1)[2\delta_2 + \eta(\delta_1)] < \epsilon$.

It is clear that a $\delta_1 > 0$ and $\delta_2 > 0$ exist so that these hold. Hence 2.3 is true.

5. *Proof of* 2.4. We are given an $\epsilon > 0$. Suppose F_ϵ is not open. Then there exists an $f \in F_\epsilon$, and for each j an f_j not in F_ϵ, such that $|f_j - f| < 1/j$. Since f_j is not in F_ϵ, one of 2.21, 2.22 fails for f_j. But 2.21 will be true for j sufficiently large, since the ϵ-mappings form an open subset[8] of F. Hence for sufficiently large j there will be an $(n + 1)$-dimensional subspace R^j of R_{2n+1} such that $f_j(M) \cdot R^j$ does not admit a finite, open ϵ-covering of order not exceeding 0. Hence R^j intersects $f_j(M)$, and therefore contains a point at a distance less than $1/j$ from the compact set $f(M)$, since $|f - f_j| < 1/j$. Thus there is a compact subset N of R_{2n+1} such that $R^j \cdot N \neq 0$ for j sufficiently large. Then there is at least one limit space R, of dimension $n + 1$, and a subsequence $j_1, j_2, \cdots$ such that the spaces $R^{j_1}, R^{j_2}, \cdots$ converge to R. To simplify the notation we now denote this subsequence as $R^1, R^2, \cdots$, and the corresponding mapping $f_{j_1}, f_{j_2}, \cdots$ as $f_1, f_2, \cdots$. We still have (1) $|f_j - f| < 1/j$ and (2) $f_j(M) \cdot R^j$ does not admit a finite, open ϵ-covering of order not exceeding 0. In addition, the $(n + 1)$-dimensional spaces R^j now converge to the $(n + 1)$-dimensional space R.

Let $\mathfrak{U}$ be a finite, open ϵ-covering of $f(M) \cdot R$ of order 0. In view of (2) above there is, for each j, a point $y_j \in f_j(M) \cdot R^j$ which is not in an element of $\mathfrak{U}$. Select $x_j \in M$ such that $f_j(x_j) = y_j$.

There is at least one point $x \in M$ such that some subsequence of $x_1, x_2, x_3, \cdots$ converges to x. We still write $x_j \to x$. Then $|f(x) - f_j(x_j)| \leqq |f(x) - f(x_j)| + |f(x_j) - f_j(x_j)|$. The first term on the right $\to 0$ as $j \to \infty$, by the continuity of f. The second term $< 1/j$, by definition of f_j. Since $f_j(x_j) \in R_j$ and $R_j \to R$, it follows that $f(x) \in R$. But then $f(x) \in f(M) \cdot R$, and hence $f(x)$ is in some open set of the collection $\mathfrak{U}$. But then we have a contradiction in the fact that this open set does not contain $f_j(x_j)$ for any j, yet $f_j(x_j) \to f(x)$ as $j \to \infty$. This contradiction implies that 2.4 is true.

This completes the proof of Theorem 1.2.

6. We apply Theorem 1.1 to give a short proof of the following theorem of Hurewicz:[9]

If M is a separable metric space of dimension n, then for any Euclidean space R_k ($k \leqq n$) there is a continuous mapping g of M into R_k such that for any p in $g(M)$ the set $g^{-1}(p)$ has dimension not exceeding $n - k$.

[8] W. Hurewicz, *Über Abbildungen von endlich dimensionalen Räumen auf Teilmengen Cartesischer Räume*, Prussian Academy of Sciences, vol. 24(1933), p. 757.

[9] Loc. cit., p. 765.

Let f be a homeomorphism of M into R_{2n+1} having the properties stated in Theorem 1.1. Let R_k be the subset of R_{2n+1} having $x_i = 0$ for $i > k$. Let h denote the orthogonal projection of R_{2n+1} into R_k; i.e., if $x = (x_1, x_2, \cdots, x_{2n+1})$, then $h(x) = (x_1, \cdots, x_k, 0, 0, \cdots, 0)$. Then set $g = hf$, and g has the desired properties. For let p be any point of $g(M)$, $p = (p_1, p_2, \cdots, p_{2n+1})$. Then the set $h^{-1}(p)$ is contained in each of the hyperplanes $x_i = p_i$ $(i = 1, 2, \cdots, k)$, hence in their intersection, a Euclidean space R of dimension $2n + 1 - k$. Then $\dim (f(M) \cdot R) \leqq n - k$, whence $\dim g^{-1}(p) \leqq n - k$.

7. Further preliminary considerations are needed before we can prove Theorem 9.1.

7.1. Definition. In a Euclidean space R_h $(h > 0)$ a collection $\mathfrak{C}$ of hyperplanes (linear subspaces of dimension $h - 1$) will be called a *Cantor collection of hyperplanes* provided that (1) the elements of $\mathfrak{C}$ form a parallel family, and (2) there is a line l whose intersections with the elements of $\mathfrak{C}$ form a closed, zero-dimensional set on l. This set may or may not be bounded.

Theorem 7.2. *Suppose that $\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_k$ are Cantor collections of hyperplanes in a Euclidean space R_h $(0 < k \leqq h)$. Let H_i be the sum of the elements of $\mathfrak{C}_i$ $(i = 1, 2, \cdots, k)$. Suppose M is a closed set in R_h such that* (1) $M \subset H_i$ *for every i, and* (2) *if for each i, P_i is any element of $\mathfrak{C}_i$ then* $\dim \left(M \cdot \prod_{i=1}^{k} P_i\right) \leqq n$. *Then* $\dim M \leqq n$.

We first prove the theorem for arbitrary h, but with $k = 1$. For simplicity we assume that the elements of $\mathfrak{C}_1$ have equations of the form $x_1 = c$. For $P \in \mathfrak{C}_1$ we have $\dim (M \cdot P) \leqq n$, and we are to show that $\dim M \leqq n$. We limit ourselves to an arbitrary compact subset I of R_h. The proof will depend upon the Lebesgue covering theorem. That is, for any $\epsilon > 0$, we show the existence of a finite, open ϵ-covering of $M \cdot I$ of order not exceeding n. This yields the desired result.

Suppose $\epsilon > 0$ is given. Select any $P \in \mathfrak{C}_1$, and suppose it has equation $x_1 = r$. Then there is a finite open ϵ-covering $\mathfrak{U}_r$ of order not exceeding n of $M \cdot I \cdot P$. Since M is closed, it follows that $\mathfrak{U}_r$ will be a covering of $M \cdot I \cdot P'$ for all elements $P' \in \mathfrak{C}_1$ which are sufficiently close to P. We can select two real numbers, r_1 and r_2, such that (1) $r_1 < r < r_2$, (2) the hyperplanes $x_1 = r_1$ and $x_2 = r_2$ are not in $\mathfrak{C}_1$, and (3) $\mathfrak{U}_r$ covers $M \cdot I \cdot P'$ if P' lies between the hyperplanes $x_1 = r_1$ and $x_1 = r_2$. Now modify the elements of the covering $\mathfrak{U}_r$ so that the modified covering still covers $M \cdot I \cdot P'$ as above, and in addition has its elements lying between the hyperplanes $x_1 = r_1$ and $x_1 = r_2$.

We obtain such a covering $\mathfrak{U}_r$ for every P which intersects I. By an application of the Borel-Lebesgue theorem to an interval of the x_1-axis we obtain a finite number of such coverings $\mathfrak{U}_r$ which together cover all of $M \cdot I$. These can then be modified so that no element of one of the coverings $\mathfrak{U}_r$ intersects an element of another covering $\mathfrak{U}_r$. Then the collection of all elements of all these

sets $\mathfrak{U}_r$ will be a finite, open ϵ-covering of $M \cdot I$ of order not exceeding n. This completes the proof for the case $k = 1$.

We now suppose the theorem true for $h < h_1$ and prove it for $h = h_1$. Let P be any element of $\mathfrak{C}_k$. Then P is a Euclidean space R_h with $h < h_1$. We can relativize the collections $\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_{k-1}$ and the sets $M, H_1, H_2, \cdots, H_{k-1}$ to the space R_h ($= P$) and apply our inductive hypothesis, yielding this result: $\dim (M \cdot P) \leqq n$. Then in the space R_{h_1} we have the one Cantor collection of hyperplanes $\mathfrak{C}_k$, with $M \subset H_k$, and for every hyperplane of the collection $\mathfrak{C}_k$ we have $\dim (M \cdot P) \leqq n$. Hence by the first result proved we conclude that $\dim M \leqq n$.

8. THEOREM 8.1. *Suppose $s_1, s_2, \cdots, s_i$ are real numbers such that if $\sum_{j=1}^{i} r_j s_j = r$, with r and every r_j rational, then every $r_j = 0$. Then there exists a 0-dimensional perfect set C of real numbers with the following properties:*

(1) *C has neither upper nor lower bound;*

(2) *the end-points of segments complementary to C are rational; and*

(3) *if $x_1, x_2, \cdots, x_i$ and $y_1, y_2, \cdots, y_i$ are in C and $\sum_{j=1}^{i} (x_j - y_j)s_j$ is rational, then $x_j = y_j$ for every j.*

Proof. Let K_0 denote the closed interval $0 \leqq x \leqq 1$. Let G_0 be the finite set of numbers -1, 0, and 1. In general, if K_n is the sum of a finite set of mutually exclusive intervals on K_0, let G_n be the finite set of numbers of the form $x - y + t$, where x and y are end-points of intervals of K_n, $|t| \leqq n$, and t is an integer. Order all non-zero rational numbers into a sequence $q_1, q_2, \cdots$.

Suppose we have defined $K_0, \cdots, K_n$ ($n \geqq 0$) so that the following properties hold:

8.2. if $k \leqq n$, K_k is the sum of 2^k mutually exclusive closed intervals having rational end-points and being of length not exceeding 2^{-k};

8.3. if $k < n$, then $K_{k+1} \subset K_k$;

8.4. if $0 < k \leqq n$, $x_1, x_2, \cdots, x_i, y_1, y_2, \cdots, y_i$ are in K_k and $t_1, t_2, \cdots, t_i$ are integers, $|t_j| < k$, then

$$\sum_{j=1}^{i} (x_j - y_j + t_j)s_j \neq q_r \qquad (0 < r \leqq k);$$

and

8.5. under hypotheses as in 8.4, if $\sum_{j=1}^{i} (x_j - y_j + t_j) = 0$, then, for every j, x_j and y_j are in the same interval of K_k.

We now define K_{n+1} so that the above properties hold with n replaced by $n + 1$. Let $F(n)$ denote the set of all linear combinations $\sum_{j=1}^{i} s_j a_j$, where $a_j \in G_n$. Now no one of these sums equals any q_k ($k \leqq n + 1$), by hypothesis on the s_j and the rationality of a_j. Let f denote any particular element of $F(n)$:

$f = \sum_{j=1}^{i} s_j a_j$. Replace a_j by z_j and write $f(z_1, z_2, \cdots, z_i) = \sum_{j=1}^{i} s_j z_j$, where the z_j are to be considered as independent real variables. Since $f(a_1, a_2, \cdots, a_i) \neq q_k$ ($k \leqq n + 1$), there is, by continuity of f, a $\delta_{fk} > 0$ such that if $| a_j - z_j | < \delta_{fk}$ for $j \leqq i$, then $f(z_1, z_2, \cdots, z_i) \neq q_k$. But since the number of elements $f \in F(n)$ is finite, there is a $\delta_1 > 0$ such that $\delta_1 < \delta_{fk}$ for every $f \in F(n)$ and every $k \leqq n + 1$.

Similarly, if some $a_j \neq 0$, then $f(a_1, a_2, \cdots, a_i) \neq 0$, and there is a $\delta_2 > 0$, independent of f, such that $f(z_1, z_2, \cdots, z_i) \neq 0$ if $| a_j - z_j | < \delta_2$ for $j \leqq i$. Let δ be the smaller of δ_1 and δ_2. To get K_{n+1} we take out from each interval I of the set of 2^n intervals whose sum is K_n an open interval I', concentric with I, having rational end-points, and such that each of the two closed intervals whose sum is $I - I'$ has length less than $\delta/2$. Then K_{n+1} is the sum of the 2^{n+1} closed intervals remaining.

Suppose now that $x_1, x_2, \cdots, x_i, y_1, y_2, \cdots, y_i$ are in K_{n+1}, and $t_1, t_2, \cdots, t_i$ are integers with absolute value less than $n + 1$. Then for each j there are points c_j and d_j which are end-points of K_n, such that $| x_j - c_j | < \delta/2$ and $| y_j - d_j | < \delta/2$. Set $a_j = c_j - d_j + t_j$, and $z_j = x_j - y_j + t_j$. Then $a_j \in G_n$ and $\sum_{j=1}^{i} (x_j - y_j + t_j)s_j = \sum_{j=1}^{i} s_j z_j = f(z_1, z_2, \cdots, z_i)$, with $f \in F(n)$. Furthermore, $| z_j - a_j | = | x_j - y_j - c_j + d_j | \leqq | x_j - c_j | + | y_j - d_j | < \delta$. It follows that $f(z_1, z_2, \cdots, z_i) \neq q_k$ for any $k \leqq n + 1$. Also $f(z_1, z_2, \cdots, z_i) \neq 0$ unless every $a_j = 0$. But in this case $c_j = d_j$, and x_j and y_j must be in the same interval of K_{n+1}. Thus the properties 8.2, 8.3, 8.4, and 8.5 hold. Let C_1 denote the common part of $K_0, K_1, \cdots$. Suppose $x_1, x_2, \cdots, x_i, y_1, y_2, \cdots, y_i$ are in C_1 and $t_1, t_2, \cdots, t_i$ are any integers. Suppose $\sum_{j=1}^{i} (x_j - y_j + t_j)s_j = r$, where r is rational. Then either $r = 0$ or $r = q_k$ for some k. Suppose $r = q_k$. Choose $n > k$ and $n > | t_j |$ ($j = 1, 2, \cdots, i$). It follows, since $C_1 \in K_n$ and $k < n$, that $r \neq q_k$. If $r = 0$, it follows that x_j and y_j are in the same interval of K_n, hence $| x_j - y_j | < 2^n$, for every n, hence $x_j = y_j$.

Now let C be the set of all numbers of the form $x + t$, where $x \in C_1$ and t is any integer. Then C has the desired properties. For if x_j and y_j are in C, then $x_j - y_j = x_j' - y_j' + t_j'$, where x_j' and y_j' are in C_1, and t_j' is an integer. This completes the proof of Theorem 8.1.

9. Theorem 9.1. *If the separable metric space M has dimension n, then there exist spaces $M_1, M_2, \cdots, M_n$ and continuous mappings $\varphi_1, \varphi_2, \cdots, \varphi_n$ such that for each j ($1 \leqq j \leqq n$) we have* (1) $\varphi_j(M_j) = M$, (2) *for $y \in M$, $\varphi_j^{-1}(y)$ consists of $j + 1$ or fewer points, and* (3) $\dim M_j = n - j$.

We derive Theorem 9.1 from the following

Theorem 9.2. *Let R_k denote k-dimensional Euclidean space. Then there exist k Cantor collections of hyperplanes $\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_k$, and continuous mappings $\varphi_1, \varphi_2, \cdots, \varphi_k$, such that if H_j is the sum of the elements of $\mathfrak{C}_j$ and $N_j = H_1 \cdot H_2 \cdot \cdots \cdot H_j$, then the following properties hold for* $1 \leqq j \leqq k$:

(1) $\varphi_j(N_j) = R_k$,

(2) if $y \in R_k$ then $\varphi_j^{-1}(y)$ consists of $j + 1$ or fewer points, and

(3) if P_{k-j} is a Euclidean $(k - j)$-space which is common to an element of each of $\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_j$, then on P_{k-j} the mapping φ_j is a homeomorphism, and $\varphi_j(P_{k-j})$ is a linear $(k - j)$-space in R_k.

Proof of Theorem 9.2. Let $s_1, s_2, \cdots, s_k$ be[10] a set of real numbers satisfying the hypothesis of Theorem 8.1 with $k = i$ and let C be a 0-dimensional perfect set of real numbers such that the conclusions of Theorem 8.1 hold. Then there is[11] a continuous mapping g of C onto R_1 (the set of all real numbers) with the following properties: (1) if $x \in C$, $y \in C$, and $x < y$, then $g(x) \leqq g(y)$, and (2) for any $p \in R_1$, $g^{-1}(p)$ consists of one or two points and if $g^{-1}(p)$ consists of two points then p is rational. For $1 \leqq j \leqq k$, let $\mathfrak{C}_j$ denote the Cantor set of hyperplanes having equations $x_j = c$, $c \in C$. Then H_j and N_j are determined, as stated in the theorem.

9.3. Definition of φ_j. Suppose $x = (x_1, \cdots, x_k)$ is in N_j. Define x_i' as follows:

9.31. if $i \leqq j$, $x_i' = g(x_i) + \sum_{r=1}^{i-1} s_r x_r$,

9.32. if $j < i \leqq k$, $x_i' = x_i$.

Then $\varphi_j(x) = (x_1', x_2', \cdots, x_k')$.

Proof that $\varphi_j(N_j) = R_k$. Note first that φ_j is actually defined for $x \in N_j$; for $x \in N_j$ implies that $x_i \in C$, $i \leqq j$, hence g is defined. Now take any $y = (y_1, \cdots, y_k) \in R_k$. We want an $x \in N_j$ such that $\varphi_j(x) = y$. Let x_1 be a value of $g^{-1}(y_1)$. Suppose $x_1, \cdots, x_{i-1}$ have been defined so that $x_r' = y_r$ $(r \leqq i - 1 < j)$. Then let x_i be a value of $g^{-1}[y_i - \sum_{r=1}^{i-1} s_r x_r]$. It follows that $x_i' = y_i$. For $i > j$, let $x_i = y_i$. For the point $x = (x_1, \cdots, x_k)$ so determined we have $\varphi_j(x) = y$.

Proof that $\varphi_j^{-1}(z)$ consists of $j + 1$ or fewer points. For any i, $1 \leqq i \leqq k$, we define a mapping f_i of N_i onto N_{i-1} (where by N_0 we mean R_k). For $x \in N_i$ let $f_i(x)$ have all coördinates the same as x except the i-th one, which shall be x_i', as given by formula 9.31. Then $\varphi_j = \varphi_{j-1}f_j = \varphi_{j-2}f_{j-1}f_j = f_1 \cdots f_j$. Now every f_i is at most 2-to-1, hence it follows that φ_j is at most 2^j-to-1. But this result is not strong enough. We prove the following.

9.4. *If $z \in R_k$, and x and y are distinct points of N_{i-1} such that $\varphi_{i-1}(x) = \varphi_{i-1}(y) = z$, then at least one of $f_i^{-1}(x)$, $f_i^{-1}(y)$ consists of a single point.*

Write $z = (z_1, z_2, \cdots, z_k)$. Then we can put $x = (x_1, x_2, \cdots, x_{i-1}, z_i, \cdots, z_k)$ and $y = (y_1, y_2, \cdots, y_{i-1}, z_i, \cdots, z_k)$, since $\varphi_{i-1}(x) = \varphi_{i-1}(y) = z$, and φ_{i-1} can change only the first $i - 1$ coördinates. Suppose both $f_i^{-1}(x)$ and

[10] For example, s_j can be taken as the square root of the j-th prime.

[11] The mapping g can be defined as follows: Order the maximal open segments of $R_1 - C$ into a countable sequence $t_1, t_2, \cdots$. Next order the rational numbers into a sequence $r_1, r_2, \cdots$ such that for every $i > 0$, r_{i+1} is in the same position relative to $r_1, r_2, \cdots, r_i$ (as regards linear order) as t_{i+1} is relative to $t_1, t_2, \cdots, t_i$. Then the stipulation that g is to take the end-points of t_i into r_i, and to be continuous, defines g.

$f_i^{-1}(y)$ consist of two points. The i-th coördinates of the points in $f_i^{-1}(x)$ and $f_i^{-1}(y)$ are respectively $g^{-1}[z_i - \sum_{r=1}^{i-1} s_r x_r]$ and $g^{-1}[z_i - \sum_{r=1}^{i-1} s_r y_r]$. By assumption that g^{-1} is 2-valued, it follows that the arguments are rational. Then their difference is rational, and we write $\sum_{r=1}^{i-1} (x_r - y_r)s_r = t$, rational. But x_r and y_r are in C $(r \leqq i - 1)$, since x and y are in N_{i-1}. It follows (see Theorem 8.1) that $x_r = y_r$; hence $x = y$. This is a contradiction. Hence of all the points of N_{i-1} which go into a point z of R_k under φ_{i-1}, there is at most one of them which has two inverses under f_i. Then $\varphi_i^{-1}(z)$ can consist of at most one point more than $\varphi_{i-1}^{-1}(z)$. Since $\varphi_1^{-1}(z)$ consists of at most two points, it follows that $\varphi_j^{-1}(z)$ consists of at most $j + 1$ points.

Now let P be a $(k - j)$-dimensional subspace of R_k common to the j hyperplanes with equations $x_1 = c_1, x_2 = c_2, \cdots, x_j = c_j$, where $c_1, c_2, \cdots, c_j$ are arbitrary elements of C. For $x \in P$ we have $x = (c_1, c_2, \cdots, c_j, x_{j+1}, \cdots, x_k)$. Then $\varphi_j(x) = (c_1', c_2', \cdots, c_j', x_{j+1}, \cdots, x_k)$, where $c_1', c_2', \cdots, c_j'$ are given by formula 9.31. Clearly φ_j is a homeomorphism on P, and $\varphi_j(P)$ is the $(k - j)$-dimensional space common to the j hyperplanes $x_1 = c_1', x_2 = c_2', \cdots, x_j = c_j'$.

9. *Proof of Theorem* 9.1. Let f be a homeomorphic mapping of M into R_{2n+1} having the properties stated in Theorem 1.1. That is, for any m-dimensional linear subspace R_m $(n + 1 \leqq m \leqq 2n + 1)$, we have $\dim (f(M)\cdot R_m) \leqq m - n - 1$. Write $M_0 = f(M)$. Apply Theorem 9.2 to the space R_{2n+1}, obtaining sets $N_1, N_2, \cdots, N_{2n+1}$ and mappings $\varphi_1, \varphi_2, \cdots, \varphi_{2n+1}$. For each j we have $N_j = \varphi_j^{-1}(R_{2n+1})$. For $1 \leqq j \leqq n$, let M_j denote $\varphi_j^{-1}(M_0)$. That is, M_j is the subset of N_j consisting of all points which go into M_0 under φ_j.

Then $\varphi_j(M_j) = M_0$, and φ_j is an at most $(j + 1)$-to-1 mapping. It remains to show that $\dim M_j = n - j$. Now it is known[12] that $\dim M_j \geqq n - j$, so we need only show that $\dim M_j \leqq n - j$. Now $M_j \subset N_j$, and N_j is common to the sums $H_1, H_2, \cdots, H_j$ of the Cantor collections $\mathfrak{C}_1, \mathfrak{C}_2, \cdots, \mathfrak{C}_j$ of hyperplanes. Hence by Theorem 7.2 it will follow that $\dim M_j \leqq n - j$, if we prove the following: If P is the common part of the j hyperplanes $x_1 = c_1, x_2 = c_2, \cdots, x_j = c_j$, where $c_1, c_2, \cdots, c_j$ are in C, then $\dim (M_j\cdot P) \leqq n - j$. Now the transformation φ_j (as defined over all N_j) takes P homeomorphically into a $(2n + 1 - j)$-dimensional linear subspace $\varphi_j(P)$ of R_{2n+1}, and $\dim (M_0\cdot\varphi_j(P)) \leqq 2n + 1 - j - n - 1 = n - j$. Then since φ_j is a homeomorphism over P, it follows that $\dim \varphi_j^{-1}(M_0\cdot\varphi_j(P)) = \dim (M_j\cdot P) = n - j$. The transformations $\varphi_1, \varphi_2, \cdots, \varphi_n$ carry the sets $M_1, M_2, \cdots, M_n$ into the set M_0 in the desired way. To get mappings into M we need only apply the inverse of the homeomorphism f.

DUKE UNIVERSITY.

[12] See footnote 1.

SIMPLE EXPLICIT EXPRESSIONS FOR GENERALIZED BERNOULLI NUMBERS OF THE FIRST ORDER

By H. S. Vandiver

Many different explicit expressions have been given for the Bernoulli numbers, and in many ways the simplest is the following, due to Kronecker:[1]

$$b_{n-1} = \sum_{a=1}^{n} \binom{n}{a} \frac{S_{n-1}(a)}{a} (-1)^{a-1}, \tag{1}$$

where

$$S_{n-1}(a) = 0^{n-1} + 1^{n-1} + 2^{n-1} + \cdots + (a-1)^{n-1}, \qquad 0^0 = 1,$$

the b's being defined by the recursion formula $(b+1)^n = b_n$, $n > 1$, where after expansion by the binomial theorem we set $b^k = b_k$.

In the present note we shall consider what is called by the writer the generalized Bernoulli number of the first order,[2]

$$(mb + k)^n = b_n(m, k), \tag{2}$$

where this is to be interpreted symbolically as in the expression involving b above, and where m and k are integers, $m \neq 0$. We have, obviously, $b_n = b_n(1, 0)$.

We shall derive explicit expressions for this generalized number which include (1) as a special case, and a number of more general forms for (1). It will be shown that these explicit expressions will yield a number of properties of the generalized Bernoulli numbers which include most of the known arithmetical properties of the ordinary Bernoulli numbers.

Our point of departure is the formula[3]

$$(b(m, k) + rm)^{n+1} - b_{n+1}(m, k) = m(n+1) \sum_{i=0}^{r-1} (im + k)^n; \tag{3}$$

another proof was given by the writer.[4] Then, in particular, the special case of this when $r = 1$, which may be written

$$(b(m, k) + m)^{n+1} - b_{n+1}(m, k) = m(n+1)k^n, \tag{4}$$

Received February 18, 1941.

[1] L. Kronecker, *Werke*, vol. 2, Leipzig, 1897, pp. 405–406.

[2] H. S. Vandiver, *On generalizations of the numbers of Bernoulli and Euler*, Proceedings of the National Academy of Sciences, vol. 23(1937), pp. 555–559.

[3] J. W. L. Glaisher, *On the value of certain series*, Quarterly Journal of Mathematics, vol. 31(1900), pp. 193–227; pp. 193–199.

[4] H. S. Vandiver, *An extension of the Bernoulli summation formula*, American Mathematical Monthly, vol. 36(1929), pp. 36–37.

may be used as a recursion formula for (2). Set

$$\sum_{i=0}^{r-1} (im + k)^n = S_n(m, k, r); \tag{5}$$

expanding (3) and setting $n - 1$ instead of n, with $n > 0$, we obtain

$$nb_{n-1}(m, k)rm + \binom{n}{2} b_{n-2}(m, k)r^2m^2 + \cdots + (rm)^n = mnS_{n-1}(m, k, r). \tag{6}$$

We shall now prove

THEOREM I. *If m and k are integers, and*

$$S_{n-1}(m, k, a) = \sum_{i=0}^{a-1} (mi + k)^{n-1},$$

then, for $r \geqq n$,

$$\sum_{a=1}^{r} \binom{r}{a}(-1)^{a-1} \frac{S_{n-1}(m, k, a)}{a} = (mb + k)^{n-1}, \tag{7}$$

where the expression on the left is to be expanded in full by the binomial theorem, and b_s set for b^s.

Proof.[5] Using (6), we may write

$$f(a) = \frac{1}{n} \sum_{i=1}^{n} \binom{n}{i} b_{n-i}(m, k)a^{i-1}m^{i-1} = \frac{S_{n-1}(m, k, a)}{a}. \tag{8}$$

Since the degree of $f(a) < n$, we have, using finite differences,

$$\Delta^r f(a) = 0, \tag{9}$$

for $r \geqq n$, and from a known formula we obtain

$$\sum_{j=1}^{r} (-1)^{j-1}\binom{r}{j} f(a + j) = f(a). \tag{10}$$

Put $a = 0$. Then

$$f(0) = b_{n-1}(m, k) = \sum_{j=1}^{r} (-1)^{j-1}\binom{r}{j} \frac{S_{n-1}(m, k, j)}{j}. \tag{11}$$

This is the result, if now we set $j = a$.

We may note also that we may extend the theorem so that m and k are quantities in any ring which contains the rational field except that m is not a zero divisor.

We now proceed to establish some other explicit forms for $b_{n-1}(m, k)$. Let m and k now be rational integers; we note that

$$e^{vn} - ne^{v(n-1)} + \cdots + (-1)^n = (e^v - 1)^n, \tag{12}$$

[5] I am indebted to the referee of this paper for the proof which follows. My own argument for the proof of the theorem was not nearly so simple.

where e is the Napierian base. This may be written

$$\sum_{a=1}^{n}\binom{n}{a}(-1)^{n-a}e^{va} = (-1)^{n-1}e^{v} + \sum_{s=1}^{n-1}(e^{v}-1)^{s}e^{v}(-1)^{n-1-s}. \tag{13}$$

Integrating each member of this with respect to v between the limits v and 0, and noting that

$$\int_0^v (e^v - 1)^s e^v\, dv = \frac{(e^v-1)^{s+1}}{s+1}, \tag{14}$$

we obtain

$$\sum_{a=1}^{n}\binom{n}{a}(-1)^{n-a}\frac{e^{va}}{a} - \sum_{a=1}^{n}\binom{n}{a}(-1)^{n-a}\frac{1}{a} \tag{15}$$
$$= (-1)^{n-1}e^{v} + (-1)^{n} + \sum_{s=1}^{n-1}\frac{(e^{v}-1)^{s+1}}{s+1}(-1)^{n-1-s},$$

or

$$\sum_{a=1}^{n}\binom{n}{a}(-1)^{n-a}\frac{e^{va}-1}{a} = (-1)^{n-1}e^{v} + (-1)^{n} + \sum_{s=1}^{n-1}\frac{(e^{v}-1)^{s+1}}{s+1}(-1)^{n-1-s}. \tag{16}$$

Dividing through by $(e^v - 1)$, we obtain

$$\sum_{a=1}^{n}\binom{n}{a}(-1)^{n-a}\frac{1}{a}\sum e^{vi} = (-1)^{n-1} + \sum_{s=1}^{n-1}\frac{(e^{v}-1)^{s}}{s+1}(-1)^{n-1-s}. \tag{17}$$

In this relation set e^m in place of e and multiply each member by e^{kv}. Now differentiate each member $n - 1$ times with respect to v and set $v = 0$. Using (7) we obtain

$$(mb + k)^{n-1} = \sum_{s=0}^{n-1}\left[\frac{d^{n-1}}{dv^{n-1}}\frac{(e^{vm}-1)^{s}e^{vk}}{s+1}\right]_{v=0}. \tag{18}$$

We then note that the right member equals

$$\sum_{s=0}^{r}\left[\frac{d^{n-1}}{dv^{n-1}}\frac{e^{vk}(e^{vm}-1)^{s}}{s+1}(-1)^{s}\right]_{v=0} \tag{19}$$

for $r \geqq n - 1$, since, using Leibnitz's theorem, we have

$$\frac{d^{n-1}}{dv^{n-1}}\left(e^{vk}(e^{vm}-1)^{s}\right) \tag{20}$$
$$= e^{vk}\frac{d^{n-1}}{dv^{n-1}}(e^{vm}-1)^{s} + k\, e^{vk}\frac{d^{n-2}}{dv^{n-2}}(e^{vm}-1)^{s}(n-1) + \cdots$$

and

$$\left[\frac{d^{h}}{dv^{h}}(e^{vm}-1)^{s}\right]_{v=0} = 0 \qquad \text{for } h < s. \tag{21}$$

Now we note that

$$\frac{d}{dv}\left(e^{vk}(e^{vm}-1)^n\right) = e^{vk}\frac{d}{dv}(e^{vm}-1)^n + (e^{vm}-1)^n k e^{vk} \tag{22}$$
$$= e^{vk}\cdot nm(e^{vm}-1)^{n-1}e^{vm} + (e^{vm}-1)^n k e^{vk}.$$

When we now take the $(n-1)$-th derivative of this with respect to v, every expression obtained by differentiating the term involving $(e^{vm}-1)^n$ will vanish on setting $v = 0$, so we examine this differentiation in connection with the first term only. We note first that, if we differentiate as a product,

$$\frac{d}{dv} e^{v(k+m)} nm(e^{vm}-1)^{n-1} = e^{v(k+m)} n(n-1)m^2(e^{vm}-1)^{n-2}e^{vm} + \cdots,$$

where the terms on the right after the first all vanish when the $(n-2)$-th derivative is taken and v made zero. We then find by induction

$$\left[\frac{d^n}{dv^n} e^{vk}(e^{vm}-1)^n\right]_{v=0} = n!\, m^n. \tag{23}$$

In (7) put S_n in place of S_{n-1} and $r = n + 1$. Also employ (17) with $(n + 1)$ in place of n; then we find

$$(mb+k)^n - \sum_{a=1}^{n}(-1)^{a-1}\frac{S_n(m,k,a)}{a}\binom{n}{a} = (-1)^n\left[\frac{d^n}{dv^n}\frac{(e^{vm}-1)^n e^{vk}}{n+1}\right]_{v=0},$$

or by (23)

$$(mb+k)^n = (-1)^n\frac{n!\,m^n}{n+1} + \sum_{a=1}^{n}(-1)^{a-1}\binom{n}{a}\frac{S_n(m,k,a)}{a}. \tag{24}$$

From the formulas (7) and (24) we shall now show how to obtain a number of results which include as special cases most of the arithmetical properties of the ordinary Bernoulli numbers. Formula (24) looks more complicated than (7) (for n in place of $(n-1)$ and $r = (n+1)$), but it does not contain any more terms and is more convenient for certain purposes. In (7) set $m = 1$, $k = 0$, and we obtain Kronecker's (1). We obtain a companion formula to this by setting $m = 1$, $k = 0$, in (24), giving

$$b_n = (-1)^n\frac{n!}{n+1} + \sum_{a=1}^{n}(-1)^{a-1}\binom{n}{a}\frac{S_n}{a}. \tag{25}$$

Let p be a prime and set $n = p$ in (24). The expression

$$(im+k)^{p-1}$$

is congruent to a, mod p, using Fermat's Theorem, unless it contains the term $mc + k$, where c is such that $mc + k \equiv 0$, mod p, in which case our expression is congruent to $a - 1$, mod p. In particular, if we set $m = 1$, $k = 0$, $r = n$ we obtain the known congruence

$$b_{p-1} \equiv 1 + \frac{(p-1)!}{p} \pmod{p}, \tag{26}$$

if we note that

$$\binom{p-1}{a} \equiv (-1)^a \qquad (\text{mod } p),$$

and

$$\sum_{a=1}^{p-1} \frac{1}{a} \equiv 0 \qquad (\text{mod } p).$$

Now consider (24) again. In the expression

$$\sum_{i=0}^{a-1} (k + im)^n \tag{27}$$

if $(m, a) = 1$, then

$$k + i_1 m \equiv k + i_2 m \qquad (\text{mod } a),$$

if and only if

$$i_1 \equiv i_2 \qquad (\text{mod } a)$$

or

$$i_1 = i_2 .$$

Hence, in (27) the quantities inside the parentheses are incongruent, modulo a, and are therefore congruent to $0, 1, \cdots, a - 1$ in some order. Hence,

$$\sum_{i=0}^{a-1} (k + im)^n \equiv \sum_{j=1}^{a-1} j^n \qquad (\text{mod } a). \tag{28}$$

It is known that if

$$n = p^\alpha n_1 , \qquad (n_1 , p) = 1, \qquad \beta \leqq \alpha,$$

then

$$\sum_{c=1}^{p^\beta h} c^{p^\alpha n_1} \equiv 0 \ (\text{mod } p^{\alpha+\beta}), \qquad n_1 \not\equiv 0 \ (\text{mod } \overline{p-1}),$$

and this may be established without the use of Bernoulli numbers. We also have[6]

$$\binom{n}{p^\beta h} \equiv 0 \qquad (\text{mod } p^{\alpha-\beta})$$

if $(h, p) = 1$; hence, in (24)

$$\frac{S_n(m, k, a)}{a} \binom{n}{a} \equiv 0 \qquad (\text{mod } p^\alpha) \qquad\qquad \text{for } n \not\equiv 0 \ (\text{mod } \overline{p-1}).$$

[6] H. S. Vandiver, *On power characters of singular integers in a properly irregular cyclotomic field*, Transactions of the American Mathematical Society, vol. 32(1930), pp. 391–408; pp. 401–402.

Now

$$\frac{n!}{n+1} \equiv 0 \pmod{p^\alpha},$$

noting that $(n + 1, p) = 1$; hence, (24) gives

THEOREM II. *If* $n = n_1 p^\alpha$, $n_1 \not\equiv 0 \pmod{p-1}$, $(m, p) = 1$, *then*

$$(mb + k)^n \equiv 0 \pmod{p^\alpha},$$

where m *and* k *are integers,* p *prime.*

The special case when $m = 1$, $k = 0$ is well known.

We shall now show that $(mb + k)^n$ is an integer for n odd, except for $n = 1$ with m odd. We have, if $(a, m) = d$, $a = da_1$,

$$\sum_{i=0}^{a-1} (k + im)^n = d \sum_{i_1=0}^{a_1-1} (k + i_1 m)^n \pmod{a}, \tag{29}$$

since if $i = a_1 r + i_1$, then

$$k + im \equiv k + i_1 m \pmod{a},$$

noting that $a_1 m \equiv 0 \pmod{a}$. Now also

$$\sum_{i_1=0}^{a_1-1} (k + i_1 m)^n \equiv \sum_{j=1}^{a_1-1} j^n \pmod{a}$$

since $(m, a_1) = 1$, so that this gives, with (29),

$$\sum_{i=0}^{a-1} (k + im)^n \equiv d \sum_{j=1}^{a_1-1} j^n \pmod{a}. \tag{30}$$

Now if a_1 is odd, then

$$\sum_{j=1}^{a_1-1} j^n = \sum_{r=1}^{(a_1-1)/2} (r^n + (a_1 - r)^n) \equiv 0 \pmod{a_1};$$

and if a_1 is even,

$$\sum_{j=1}^{a_1-1} j^n = \sum_{h=1}^{a_1/2-1} \left(h^n + (a_1 - h)^n + \left(\frac{a_1}{2}\right)^n\right) \equiv 0 \pmod{a_1}$$

for $n > 1$. Hence (30) gives

$$\sum_{i=1}^{a-1} (k + im)^n \equiv 0 \pmod{a}$$

for $n > 1$ and (7) shows that $(mb + k)^n$ is an integer for $n > 1$. For $n = 1$, $mb + k$ is obviously an integer except for m odd, and this is the result desired.

We now examine what prime factors appear in the denominators of $(mb + k)^n$, n even. The relation (30) gives

$$\frac{S_n(m, k, a)}{a} = \frac{S_n(a_1)}{a_1} + I, \tag{31}$$

where I is an integer.

By a known result[7]

$$\frac{S_n(a_1)}{a_1} + \frac{1}{2} + \frac{1}{\alpha} + \frac{1}{\beta} + \cdots$$

is an integer, where $2, \alpha, \beta, \cdots$ are the distinct von Staudt-Clausen primes corresponding to n, which also divide a_1. (We define the von Staudt-Clausen primes corresponding to n as those primes p such that $n \equiv 0 \pmod{p-1}$.) Applying this to (7) with n in place of r, using (31), and collecting the terms involving α, we find the sum to be

$$\text{(31a)} \qquad \frac{(-1)^\alpha}{\alpha}\left(\binom{n}{\alpha} - \binom{n}{2\alpha} + \binom{n}{3\alpha} - \cdots\right);$$

and employing the relation of Lucas,[8]

$$\binom{n}{m} \equiv \binom{n_1}{m_1}\binom{\nu}{\mu} \pmod{p},$$

where $m = pm_1 + \mu$, $n = pn_1 + \nu$, and p is a prime, we obtain from (31a), setting $p = \alpha$,

$$\begin{aligned} &\left(\binom{n}{\alpha} - \binom{n}{2\alpha} + \binom{n}{3\alpha} - \cdots\right) \\ &\qquad \equiv \left(\binom{n_1}{1} - \binom{n_1}{2} + \binom{n_1}{3} - \cdots\right) \pmod{\alpha} \\ &\qquad \equiv +1 \pmod{\alpha}. \end{aligned}$$

We have a similar result for each von Staudt-Clausen prime corresponding to n, and which is prime to m, since a_1 is prime to m, so that we have

THEOREM III. *We have for n even, m and k integers with $m \neq 0$,*

$$b_n(m, k) = A_n - \sum_{i=1}^{s} \frac{1}{p_i},$$

where the p's are the distinct primes which are prime to m and such that $n \equiv 0 \pmod{p_i - 1}$; A_n being some integer. For n odd, $b_n(m, k)$ is an integer, except for $n = 1$ with m odd.

This theorem was given in a previous article by the writer[9] in which an entirely different proof was indicated briefly. Many other congruences may be derived involving $b_n(m, k)$, but some of them can be more conveniently obtained from (26) instead of from (7) or (24).

We now consider various explicit expressions for $b_n(m, k)$, or the special case

[7] P. Bachmann, *Niedere Zahlentheorie*, vol. 2, Leipzig, 1910, p. 48.

[8] E. Lucas, *Théorie des fonctions numériques simplement périodiques*, American Journal of Mathematics, vol. 1(1878), pp. 184–238; pp. 229, 230.

[9] Vandiver, op. cit., footnote 2, p. 555.

of the ordinary Bernoulli number. The equation (7) leads immediately to the following:

$$\frac{(mb+k)^{n-1}}{r} = \sum_{a=1}^{r} \binom{r-1}{a-1} \frac{S_{n-1}(m, k, a)}{a^2} \tag{32}$$

with a similar relation for (24).

Now in (7) set $m = 1$; we obtain

$$\sum_{a=1}^{r} \binom{r}{a} (-1)^{a-1} \frac{\sum_{i=k}^{k+a-1} i^{n-1}}{a} = (b+k)^{n-1}. \tag{33}$$

Comparing this with the formula

$$(b+k)^{n-1} - b_{n-1} = (n-1)(1^{n-2} + \cdots + (k-1)^{n-2}), \tag{34}$$

we obtain, for $r > n$,

$$b_n = -(n-1)S_{n-1}(k) + \sum_{a=1}^{r} \binom{r}{a} (-1)^{a-1} \frac{\sum_{i=k}^{k+a-1} i^n}{a}. \tag{35}$$

This illustrates the main reason why the writer introduced generalized Bernoulli numbers, namely, to throw light on the properties of the ordinary Bernoulli numbers.

Consider the expression (18) and expand each parenthesis in the bracket on the right side of the equation and then carry out the differentiation indicated. We find after setting $v = 0$

$$(mb+k)^n = \sum_{s=0}^{n} \sum_{a=0}^{s} \frac{\binom{s}{a}(am+k)^n}{s+1} (-1)^a. \tag{36}$$

Set $m = 1$, $k = 0$, and we obtain[10]

$$b_n = \sum_{s=1}^{n} \sum_{a=0}^{s} \binom{s}{a} \frac{a^n}{s+1} (-1)^a. \tag{37}$$

We now note that (21) gives, for $m = 1$, $h < s$,

$$\sum_{a=0}^{s} (-1)^a a^h \binom{s}{a} = 0.$$

This shows that (1) may also be written

$$b_{n-1} = \sum_{a=1}^{n} \binom{n}{a} \frac{S_{n-1}(a+1)}{a} (-1)^{a-1}. \tag{38}$$

[10] J. Worpitzky, *Studien über die Bernoullischen und Eulerschen Zahlen*, Journal für Mathematik, vol. 94(1883), pp. 203–232.

This shows that in (17) we may multiply through by e^v and obtain, when we differentiate through $(n - 1)$ times with respect to v, the right member of (38) after substituting $v = 0$. This gives, by using the right member of (17), the formula

$$\sum_{s=0}^{n} \sum_{a=0}^{s} (-1)^a \frac{(a+1)^n}{s+1} \binom{s}{a} = b_n,$$

which is the Lucas formula, usually written in the form

$$b_n = 1 - \frac{\Delta^{(1)}}{2} + \frac{\Delta^{(2)}}{3} - \cdots + \frac{\Delta^{(n)}}{n+1},$$

where $\Delta^{(k)}$ represents $\Delta^{(k)}(m^n)$ or the k-th difference of m^n.

In the proof of the above we noted that

$$\left[\frac{d^n}{dv^n}\left(e^v \sum_{s=1}^{n} \frac{(1-e^v)^s}{s+1}\right)\right]_{v=0} = \left[\frac{d^n}{dv^n}\left(\sum_{s=1}^{n} \frac{(1-e^v)^s}{s+1}\right)\right]_{v=0}. \tag{39}$$

Carrying out the differentiation of the left member as a product and using Leibnitz's formula, we obtain

$$(b+1)^n = b_n. \tag{40}$$

This suggests that we may start with one of our explicit representations as a definition of a Bernoulli number and obtain the principal results in the theory of Bernoulli numbers. Following this scheme we may obtain the Bernoulli summation formula by deriving (35) independently of (34) by examining the terms in the second summation on the right of (35) noting that

$$\frac{(a+c)^{n-1}}{a} = a^{n-2} + \cdots + c^{n-2}(n-1) + \frac{c^{n-1}}{a}. \tag{41}$$

We then note also that the whole summation on the right of (35) may be obtained in the form

$$\left[\frac{d^{n-1}}{dv^{n-1}}\left(\sum_{a=1}^{n} \binom{n}{a} (-1)^{a-1} \frac{e^{(a+k)v} - e^{kv}}{e^v - 1}\right)\right]_{v=0}. \tag{42}$$

By taking the fractional expression in the above in the form

$$e^{kv} \frac{e^{av} - 1}{e^v - 1}$$

and differentiating it as a product, we may write (42) as

$$(b+k)^{n-1},$$

and this gives (34). Formula (3) may be obtained from (34) with $n + 2$ in place of n as follows, since (3) holds for $n = 0$. Assuming (3) to be true for n in place of $n + 1$, we obtain from this assumption

$$\int_0^k ((mb + k + rm)^n - (mb + k)^n)\, dk = mn \int_0^k \sum_{i=0}^{a-1} (im + k)^{n-1} dk, \tag{43}$$

or

$$\frac{(mb + k + rm)^{n+1}}{n+1} - \frac{(mb + k)^{n+1}}{n+1} - m^{n+1}\frac{((b + r)^{n+1} - b_{n+1})}{n+1}$$

$$= m \sum_{i=0}^{r-1} (im + k)^n - m^{n+1} \sum_{i=0}^{r-1} i^n ;$$

using (34) we obtain (3) for $n + 1$ in place of n, whence (3) follows in general by induction.

UNIVERSITY OF TEXAS.

GENERALIZED BERNOULLI AND EULER NUMBERS

By L. Carlitz

1. Several years ago Vandiver[1] introduced certain numbers which he called generalized Bernoulli numbers of the r-th order. In particular for $r = 1$, the generalized number of the first order was defined by means of

$$b_n(m, k) = (mb + k)^n, \tag{1.1}$$

where the right member is to be expanded and b^i replaced by b_i, the ordinary Bernoulli number:

$$(b + 1)^n = b^n \qquad (n > 1); \tag{1.2}$$

the m and k are arbitrary integers. In his paper Vandiver gave a theorem about $b_n(m, k)$ which reduces to the familiar Staudt-Clausen theorem when $m = 1$, $k = 0$; he has recently given another proof of this theorem.[2]

In the present note we shall first prove this theorem by means of Lucas'[3] method. In the remainder of the paper we consider certain Bernoulli polynomials in several variables and by the same method derive a theorem of the Staudt-Clausen type. For the corresponding Euler polynomials we derive congruences of Kummer's type; the method is that used by Nielsen[4] for the ordinary Euler numbers.

2. We require the well-known formula

$$b_n = \sum_{s=0}^{n} \frac{1}{s+1} \sum_{\alpha=0}^{s} (-1)^\alpha \binom{s}{\alpha} \alpha^n. \tag{2.1}$$

Expanding the right member of (1.1) and using (2.1) we get after some manipulation

$$b_n(m, k) = \sum_{s=0}^{n} \frac{1}{s+1} \Delta^s, \tag{2.2}$$

where for brevity we put

$$\Delta^s = \Delta^s(m, k) = \sum_{\alpha=0}^{s} (-1)^\alpha \binom{s}{\alpha} (m\alpha + k)^n. \tag{2.3}$$

Received July 11, 1941.

[1] H. S. Vandiver, *On generalizations of the numbers of Bernoulli and Euler*, Proceedings of the National Academy of Sciences, vol. 23(1937), pp. 555–559; especially p. 555.

[2] H. S. Vandiver, *Simple explicit expressions for generalized Bernoulli numbers of the first order*, this Journal, vol. 8(1941), pp. 575–584.

[3] E. Lucas, *Théorie des Nombres*, vol. 1, Paris, 1891, p. 433.

[4] N. Nielsen, *Traité Elémentaire des Nombres de Bernoulli*, Paris, 1923, p. 262.

Since

$$s! \mid \Delta^s(m, k),$$

for all s, m, k, it follows that all terms on the right of (2.2) are integral except possibly those for which $s + 1 = 4$ or $s + 1 = p$, a prime. If $s + 1 = p$, there are two possibilities:

(i) for $p \mid m$,

$$\Delta^s(m, k) \equiv \sum_{\alpha=0}^{p-1} k^n \equiv 0 \pmod{p},$$

as follows immediately from (2.3);

(ii) for $p \nmid m$,

$$\Delta^s(m, k) \equiv \sum_{\alpha=0}^{p-1} (m\alpha + k)^n \equiv \sum_{\alpha=0}^{p-1} \alpha^n \pmod{p}$$

$$\equiv \begin{cases} -1 & \text{for } p - 1 \mid n, \\ 0 & \text{otherwise.} \end{cases}$$

If $s = 3$, we have by (2.3)

$$\Delta^3 \equiv k^n + (k + m)^n - (k + 2m)^n - (k + 3m)^n \pmod{4}. \tag{2.4}$$

We consider separately three cases:

(iii) for m even,

$$\Delta^3 \equiv 0 \pmod{4},$$

as follows from (2.4) by taking $m \equiv 0, 2 \pmod{4}$;

(iv) for m odd, n even,

$$\Delta^3 \equiv (k + m)^n - (k - m)^n \equiv 0 \pmod{4};$$

(v) for m odd, n odd, we take

$$\Delta^1 + \frac{\Delta^3}{2} \equiv \frac{1}{2} \sum_{\alpha=0}^{3} (k + \alpha m)^n \pmod{2}$$

$$\equiv \begin{cases} 1 & \text{for } n = 1, \\ 0 & \text{for } n > 1. \end{cases}$$

As a consequence we get Vandiver's theorem:

For n even,

$$b_n(m, k) = G - \sum_p \frac{1}{p} \qquad (p - 1 \mid n,\ p \nmid m),$$

where G is an integer, and the summation is over all primes p satisfying the conditions indicated; for n odd, $b_n(m, k)$ is an integer, except for $n = 1$, m odd, in which case $b_1(m, k) = G + \frac{1}{2}$.

3. We now consider certain Bernoulli polynomials in several indeterminates.[5] Put

$$(3.1) \qquad B_n(x) = B_{n_1 \cdots n_t}(x_1 \cdots x_t) = (b + x_1)^{n_1} \cdots (b + x_t)^{n_t},$$

where the right member is to be expanded in full, and b^i replaced by b_i, defined in (1.2). These polynomials have various properties generalizing those of the ordinary Bernoulli polynomials. Here we are interested in certain arithmetic properties. Following (1.1) we define

$$\begin{aligned} b_{n_1 \cdots n_t} &= b_{n_1 \cdots n_t}(m_1 \cdots m_t, k_1 \cdots k_t) \\ &= (m_1 b + k_1)^{n_1} \cdots (m_t b + k_t)^{n_t} \\ &= m_1^{n_1} \cdots m_t^{n_t} B_{n_1 \cdots n_t}\left(\frac{k_1}{m_1} \cdots \frac{k_t}{m_t}\right). \end{aligned}$$

Then using (2.1) we get

$$(3.3) \qquad b_{n_1 \cdots n_t} = \sum_{s=0}^{n_1 + \cdots + n_t} \frac{1}{s+1} \Delta^s,$$

where now

$$(3.4) \qquad \Delta^s = \sum_{\alpha=0}^{s} (-1)^\alpha \binom{s}{\alpha} (m_1\alpha + k_1)^{n_1} \cdots (m_t\alpha + k_t)^{n_t}.$$

Now, as in the previous case, for arbitrary integral m_i, k_i, we have $s! \mid \Delta^s$, and therefore all terms on the right of (3.4) are integral except possibly those for which $s + 1 = 4$ or a prime p. The remainder of the discussion of §2 becomes rather complicated in the present case, and we shall state a somewhat weakened Staudt-Clausen theorem:

For integral m_i, k_i we have

$$(3.5) \qquad b_{n_1 \cdots n_t} = G + \sum_p \frac{\epsilon_p}{p} + \frac{\epsilon_4}{4},$$

where G is integral, p runs through the primes not exceeding $n_1 + \cdots + n_t$,

$$(3.6) \qquad \epsilon_p \equiv \sum_{\alpha=0}^{p-1} (m_1\alpha + k_1)^{n_1} \cdots (m_t\alpha + k_t)^{n_t} \pmod{p},$$

and $\epsilon_4 \equiv \Delta^3 \pmod 4$ and is even.

We remark that in the same way if $f(x)$ is an arbitrary polynomial of degree n with integral coefficients then (2.1) yields

$$f(b) = \sum_{s=0}^{n} \frac{1}{s+1} \sum_{\alpha=0}^{s} (-1)^\alpha \binom{s}{\alpha} f(\alpha),$$

[5] See L. Carlitz, *On arrays of numbers*, American Journal of Mathematics, vol. 54(1932), p. 751.

and therefore we get the result:

$$f(b) = G + \sum_{p \leqq n} \frac{e_p}{p} + \frac{e_4}{4},$$

where

$$e_p \equiv \sum_{\alpha=0}^{p-1} f(\alpha) \pmod{p}$$

and

$$e_4 \equiv f(0) + f(1) - f(2) - f(3) \pmod{4},$$

so that e_4 is even.

Returning to (3.5), take $t = 2$. Let

$$n_i \equiv \nu_i \pmod{p-1} \qquad (1 \leqq \nu_i \leqq p-1),$$

and put $\delta = m_1 k_2 - m_2 k_1$. Assume $p \nmid m_1 m_2$. Then for $p \nmid \delta$, it is easily verified that (3.6) implies

$$\epsilon_p \equiv \begin{cases} -2 & \text{for } \nu_1 + \nu_2 = 2p - 2, \\ -m_1^{n_1} m_2^{n_2} \left(\dfrac{\delta}{m_1 m_2}\right)^{n_1+n_2} & \text{for } p - 1 \leqq \nu_1 + \nu_2 < 2p - 2, \\ 0 & \text{for } \nu_1 + \nu_2 < p - 1; \end{cases}$$

for $p \mid \delta$, we have

$$\epsilon_p \equiv \begin{cases} -m_1^{-n_2} m_2^{-n_1} & \text{for } p - 1 \mid n_1 + n_2, \\ 0 & \text{otherwise.} \end{cases}$$

The corresponding results in the case $p \mid m_1 m_2$ may be obtained without difficulty.

4. Parallel to (3.1) we may also define certain Euler polynomials in several indeterminates. For brevity we take as definition

$$E_{n_1 \cdots n_t}(x_1 \cdots x_t) = \sum_{s=0}^{n_1+\cdots+n_t} \frac{1}{2^s} \sum_{\alpha=0}^{s} (-1)^\alpha \binom{s}{\alpha} (\alpha + x_1)^{n_1} \cdots (\alpha + x_t)^{n_t}. \tag{4.1}$$

If now we put

$$e_n = e_{n_1 \cdots n_t} = m_1^{n_1} \cdots m_t^{n_t} E_{n_1 \cdots n_t}\left(\frac{k_1}{m_1} \cdots \frac{k_t}{m_t}\right),$$

then (4.1) yields

$$e_{n_1 \cdots n_t} = \sum_{s=0}^{n_1+\cdots+n_t} \frac{1}{2^s} \Delta^s, \tag{4.2}$$

where Δ^s is defined by (3.4). Clearly

$$2^{n_1+\cdots+n_t} e_{n_1 \cdots n_t}$$

is integral for all integral m_i, k_i. Now let p be an *odd* prime. We consider the number ($\nu_i \geqq 0$)

$$\Omega_{n_1 \cdots n_t}^{\nu_1 \cdots \nu_t} = \sum_{(i)} (-1)^{i_1+\cdots+i_t} \binom{\nu_1}{i_1} \cdots \binom{\nu_t}{i_t} e_{n+i(p-1)}. \tag{4.3}$$

By (3.4) the right member becomes

$$\sum_{\alpha=0}^{s} (-1)^{\alpha} \binom{s}{\alpha} \prod_{i=1}^{t} (m_i\alpha + k_i)^{n_i}(1 - (m_i\alpha + k_i)^{p-1})^{\nu_i}. \tag{4.4}$$

Since by Fermat's theorem, for $\nu \geqq n$,

$$m^n(1 - m^{p-1})^{\nu} \equiv 0 \pmod{p^n},$$

it follows at once from (4.4) that

$$\Omega_{n_1 \cdots n_t}^{\nu_1 \cdots \nu_t} \equiv 0 \pmod{p^{n_1+\cdots+n_t}}, \tag{4.5}$$

provided $\nu_i \geqq n_i$. We may generalize (4.5) somewhat:

$$\Omega_{n_1 \cdots n_t}^{\nu_1 \cdots \nu_t} \equiv 0 \pmod{p^{\mu_1+\cdots+\mu_t}},$$

where

$$\mu_i = \min(n_i, \nu_i) \qquad (i = 1, \cdots, t).$$

DUKE UNIVERSITY.

AN APPLICATION OF THE CLASSICAL ORTHOGONAL POLYNOMIALS TO THE THEORY OF INTERPOLATION

By H. N. Laden

1. Introduction. Let

$$(1)\qquad \begin{array}{cccc} & x_{11} & & \\ & x_{12} & x_{22} & \\ \cdots & \cdots & \cdots & \cdots \\ x_{1,n} & x_{2,n} & \cdots & x_{n,n} \\ \cdots & \cdots & \cdots & \cdots \end{array} \qquad (a \leqq x_{1,n} < x_{2,n} < \cdots < x_{n,n} \leqq b)$$

be a triangular matrix of abscissas on an interval $[a, b]$. Set $\omega_n(x) = c(x - x_{1,n})(x - x_{2,n}) \cdots (x - x_{n,n})$, with c an arbitrary non-zero constant. Then,

$$(2)\qquad l_{k,n}(x) = \frac{\omega_n(x)}{(x - x_{k,n})\omega_n'(x_{k,n})}, \quad \text{with} \quad l_{k,n}(x_{j,n}) = \delta_{kj} \qquad (k, j = 1, 2, \cdots, n;\ n = 1, 2, \cdots),$$

are the fundamental polynomials of Lagrange interpolation of degree not exceeding $n - 1$ corresponding to the n-th set of abscissas. Let

$$(3)\qquad \begin{array}{cccc} & y_{11} & & \\ & y_{12} & y_{22} & \\ \cdots & \cdots & \cdots & \cdots \\ y_{1,n} & y_{2,n} & \cdots & y_{n,n} \\ \cdots & \cdots & \cdots & \cdots \end{array}$$

be a corresponding matrix of ordinates. Then, the Lagrange interpolation polynomial $L_n(x)$ has the property that

$$(4)\qquad L_n(x) = \sum_{k=1}^{n} y_{k,n} l_{k,n}(x), \qquad L_n(x_{j,n}) = y_{j,n} \qquad (j = 1, 2, \cdots, n;\ n = 1, 2, \cdots).$$

In particular, if $f(x)$ is a function defined on the finite interval $[a, b]$ and if we select $y_{k,n} = f(x_{k,n})$ ($k = 1, 2, \cdots, n$; $n = 1, 2, \cdots$), we write $L_n[f; x] \equiv L_n[f]$. The question whether $L_n[f] \to f(x)$ uniformly on $[a, b]$ as $n \to \infty$ for an arbitrary continuous $f(x)$ was settled definitively by Faber,[1] who showed that,

Received March 13, 1941. The author wishes to express his deep gratitude to Professor J. A. Shohat for his constant encouragement and valuable advice.

[1] Jahresbericht der deutschen Mathematiker-Vereinigung, vol. 23(1914), pp. 192–210.

corresponding to every choice for the matrix of abscissas, there exists a continuous function $f(x)$ on $[a, b]$ such that $L_n[f]$ does not converge to $f(x)$ uniformly on $[a, b]$.

The introduction of the Hermite interpolation polynomial (HIP) of degree $2n - 1$,

$$
(5) \qquad \begin{aligned}
&H_n(x) = \sum_{k=1}^{n} y_{k,n}\left[1 - \frac{\omega_n''(x_{k,n})}{\omega_n'(x_{k,n})}(x - x_{k,n})\right] l_{k,n}^2(x),\\
&H_n(x_{k,n}) = y_{k,n}, \qquad H_n'(x_{k,n}) = 0 \qquad (k = 1, 2, \cdots, n; n = 1, 2, \cdots),
\end{aligned}
$$

is more fruitful. Again, we write $H_n[f]$, for the n-th HIP corresponding to $f(x)$. For this interpolatory procedure, there exist matrices of abscissas such that $H_n[f] \to f(x)$ uniformly on $[a, b]$ as $n \to \infty$ for any arbitrary $f(x)$ continuous on $[a, b]$. (Szegö [7, 8], Shohat [5].[2])

A more extended interpolatory procedure is discussed by Kryloff and Stayermann [3]. In this is employed an interpolation polynomial of degree $4n - 1$ corresponding to (1) and (3):

$$
(6) \qquad F_n(x) = \sum_{k=1}^{n} y_{k,n}\, u_{k,n}(x) l_{k,n}^4(x) \qquad (n = 1, 2, \cdots),
$$

$$
(7) \qquad \begin{aligned}
u_{k,n}(x) = 1 &- 2(x - x_{k,n}) \frac{\omega_n''(x_{k,n})}{\omega_n'(x_{k,n})}\\
&+ \frac{(x - x_{k,n})^2}{2}\left\{5\left[\frac{\omega_n''(x_{k,n})}{\omega_n'(x_{k,n})}\right]^2 - \frac{4}{3}\frac{\omega_n'''(x_{k,n})}{\omega_n'(x_{k,n})}\right\}\\
&+ \frac{(x - x_{k,n})^3}{6}\left\{-15\left[\frac{\omega_n''(x_{k,n})}{\omega_n'(x_{k,n})}\right]^3 + 10\frac{\omega_n''(x_{k,n})\omega_n'''(x_{k,n})}{[\omega_n'(x_{k,n})]^2} - \frac{\omega_n^{\mathrm{IV}}(x_{k,n})}{\omega_n'(x_{k,n})}\right\},
\end{aligned}
$$

$$
(8) \qquad \begin{aligned}
&F_n(x_{k,n}) = y_{k,n}, \qquad F_n^{(\nu)}(x_{k,n}) = 0\\
&\qquad\qquad (\nu = 1, 2, 3; k = 1, 2, \cdots, n; n = 1, 2, \cdots).
\end{aligned}
$$

As before, for a function $f(x)$ on the interval $[a, b]$ for which $f(x_{k,n}) = y_{k,n}$, we write $F_n[f; x] \equiv F_n[f]$.

Kryloff and Stayermann [3] limit their considerations to the case where the n-th set of abscissas are the zeros of the trigonometric ("Tchebycheff") polynomial

$$
T_n(x) = \cos(n \arccos x) = 2^{n-1}x^n + \cdots \qquad (n = 1, 2, \cdots).
$$

They assert that for such a choice of abscissas and for an arbitrary function $f(x)$ continuous on $[-1, 1]$, $F_n[f] \to f(x)$ uniformly on $[-1, 1]$ as $n \to \infty$. Their derivation, however, is in error and we give a corrected proof. Furthermore, we consider the more general case of abscissas placed at the zeros of the classical orthogonal polynomials of Jacobi, Laguerre and Hermite, employing in the main

[2] Numbers in brackets refer to the bibliography at the end of the paper.

a method originated by Fejér [1] and extended by Shohat [5] and Szegö [7, 8]. The results in this direction lead to independent and constructive proofs of the Weierstrass approximation theorem for continuous functions. Furthermore, it is revealed that there are abscissas which produce the desired convergence theorem when employed in the HIF, which do not produce this result when used in the formula of Kryloff and Stayermann. This last may prove surprising to those expecting special advantages in this direction to accrue from the increase in the degree of the n-th interpolation polynomial.

2. **Basic convergence theorem.** Let $f(x)$ be a uniformly continuous function on an interval $[a, b]$ (finite), let (1) be a matrix of abscissas on this interval, and consider (6) for this function. Given $\epsilon > 0$ there exists a positive δ such that $|f(x') - f(x'')| < \epsilon$ for $|x' - x''| \leqq \delta$, $a \leqq x', x'' \leqq b$. Also, let $|f(x)| \leqq M$ for $a \leqq x \leqq b$. We remark that

$$\sum_{k=1}^{n} u_{k,n}(x) l_{k,n}^4(x) \equiv 1. \tag{9}$$

Thus, for a fixed x in $[a, b]$,

$$\begin{aligned} f(x) - F_n[f] &= \sum_{k=1}^{n} [f(x) - f(x_{k,n})] u_{k,n}(x) l_{k,n}^4(x), \\ |f(x) - F_n[f]| &\leqq \sum_{|x - x_{k,n}| \leqq \delta} |f(x) - f(x_{k,n})| \, | u_{k,n}(x) | \, l_{k,n}^4(x) \\ &\qquad + \sum_{|x - x_{k,n}| > \delta} |f(x) - f(x_{k,n})| \, | u_{k,n}(x) | \, l_{k,n}^4(x) \\ &\leqq \epsilon \sum_{|x - x_{k,n}| \leqq \delta} | u_{k,n}(x) | \, l_{k,n}^4(x) \\ &\qquad + 2M \sum_{|x - x_{k,n}| > \delta} | u_{k,n}(x) | \, l_{k,n}^4(x). \end{aligned} \tag{10}$$

Consequently, we have

THEOREM A. *Let $f(x)$ be an arbitrary function continuous on $[a, b]$. If, for any system of abscissas on $[a, b]$, we have both*

$$\lim_{n \to \infty} \sum_{|x - x_{k,n}| > \delta} | u_{k,n}(x) | \, l_{k,n}^4(x) = 0, \tag{11}$$

$$\sum_{|x - x_{k,n}| \leqq \delta} | u_{k,n}(x) | \, l_{k,n}^4(x) \text{ is bounded} \tag{12}$$

holding uniformly for x in any subinterval $[c, d]$ of $[a, b]$, then $F_n[f] \to f(x)$ uniformly on $[c, d]$ as $n \to \infty$.

§3 is now devoted to establishing the conditions of this theorem for trigonometric abscissas and the correction of the error of Kryloff and Stayermann. §4 establishes the theorem for a wide class of Jacobi abscissas. The remaining sections are concerned with the Laguerre and Hermite abscissas, for which the above theorem requires some modification.

3. Trigonometric abscissas. Making use of the differential equation

$$(1 - x^2)\omega_n''(x) - x\omega_n'(x) + n^2\omega_n(x) = 0, \qquad \omega_n(x) \equiv T_n(x), \tag{13}$$

and its (twice) derived equations, we get[3]

$$\begin{gathered} \frac{\omega_n''(x_k)}{\omega_n'(x_k)} = \frac{x_k}{1 - x_k^2}, \qquad \frac{\omega_n'''(x_k)}{\omega_n'(x_k)} = \frac{3x_k^2}{(1 - x_k^2)^2} - \frac{n^2 - 1}{1 - x_k^2}, \\ \frac{\omega_n^{\mathrm{IV}}(x_k)}{\omega_n'(x_k)} = \frac{15x_k^3}{(1 - x_k^2)^3} - \frac{(6n^2 - 9)x_k}{(1 - x_k^2)^2}; \end{gathered} \tag{14}$$

whence, by (7),

$$u_k(x) = \frac{1 - 2xx_k + x_k^2}{1 - x_k^2} + \frac{(x - x_k)^2}{2}\left[\frac{x_k^2 + \frac{1}{3}(n^2 - 1)(1 - x_k^2)}{(1 - x_k^2)^2}\right] + \frac{(x - x_k)^3}{6}\left[\frac{-(4n^2 - 1)x_k}{(1 - x_k^2)^2}\right].$$

Essential in the proof of Kryloff and Stayermann is the non-negativeness of $u_{k,n}(x)$ for $-1 \leqq x \leqq 1$ ($k = 1, 2, \cdots, n$; $n = 1, 2, \cdots$). This they attempt to establish in the following manner. They write

$$u_k(x) = P + Q, \qquad P \equiv \frac{1 - 2xx_k + x_k^2}{1 - x_k^2} = \frac{1 - x^2 + (x - x_k)^2}{1 - x_k^2} \geqq 0 \quad \text{for } -1 \leqq x \leqq 1,$$

$$\begin{aligned} Q &= \frac{(x - x_k)^2}{6(1 - x_k^2)^2}\{6x_k^2 - x_k[3x + 4(n^2 - 1)x] + 4(n^2 - 1)\} \\ &= \frac{(x - x_k)^2}{6(1 - x_k^2)^2}\,q(x) \end{aligned} \tag{15}$$

and proceed to consider $q(x)$, the quadratic form in x_k, for $-1 \leqq x \leqq 1$. Its discriminant is actually $x^2[3 + 4(n^2 - 1)]^2 - 96(n^2 - 1)$, although Kryloff and Stayermann write erroneously $x^2[2 + 4(n^2 - 1)]^2 - 96(n^2 - 1)^2$. From the non-positiveness of the latter, they deduce that $q(x)$ is non-negative for $-1 \leqq x \leqq 1$, provided $n \geqq 2$. However, this result cannot be deduced from the corrected discriminant, since for $n = 3$ the interval of its non-positiveness is already $\left[-\frac{16\cdot 3^{\frac{1}{2}}}{35}, \frac{16\cdot 3^{\frac{1}{2}}}{35}\right]$ and this subinterval of $[-1, 1]$ continues to shrink with increasing n.

Still the final result of Kryloff and Stayermann is correct, for

$$q(x) = 4(n^2 - \tfrac{7}{4})(1 - xx_k) + 3(1 - xx_k)^2 + 3x_k^2(1 - x^2) + 3x_k^2 \geqq 0,$$

$$Q \geqq 0, \quad u_{k,n}(x) \geqq 0 \qquad \text{for } -1 \leqq x \leqq 1 \ (k = 1, 2, \cdots, n; n \geqq 2).$$

[3] Unless there is danger of confusion, we write x_k, $\omega(x)$, $u_k(x)$, $\ldots$ in place of $x_{k,n}$, $\omega_n(x)$, $u_{k,n}(x)$, $\ldots$ respectively.

This, once established, enables us to obtain (12) and (11) uniformly in $[-1, 1]$. This makes Theorem A applicable and we have

THEOREM I (Kryloff and Stayermann). *Let $f(x)$ be an arbitrary continuous function on $[-1, 1]$. For trigonometric abscissas, $F_n[f] \to f(x)$ uniformly on $[-1, 1]$ as $n \to \infty$.*

We have omitted details concerning (11) and (12) for trigonometric abscissas, these latter being but a particular case of Jacobi abscissas, which are treated below.

4. **General Jacobi abscissas** $(\alpha, \beta > -1)$ (notations those of Szegö [5]). With two parameters, the computation in places becomes rather involved and is, of necessity, omitted.

We again use the differential equation for the Jacobi polynomial:

$$
\begin{aligned}
(16)\qquad &(1 - x^2)\omega''(x) + [\beta - \alpha - (\alpha + \beta + 2)x]\omega'(x) \\
&\qquad + n(n + \alpha + \beta + 1)\omega(x) = 0, \qquad \omega_n(x) = P_n^{(\alpha,\beta)}(x),
\end{aligned}
$$

and its (twice) derived equations as in §3, as well as (7), to get

$$
(17)\quad \left\{
\begin{aligned}
u_k(x) &= p(x) + \frac{(x - x_k)^2}{6(1 - x_k^2)^2}\, t(x), \\
p(x) &= 1 - \frac{2(x - x_k)(\alpha + 1)}{1 - x_k} + \frac{2(x - x_k)(\beta + 1)}{1 + x_k} \equiv \frac{s(x)}{1 - x_k^2}, \\
t(x) &= 4n(n + \alpha + \beta + 1)\mu_1(x) + \mu_2(x), \\
\mu_1(x) &= 1 - x_k^2 - (x - x_k)[2\alpha - 2\beta + (2\alpha + 2\beta + 3)x_k], \\
\mu_2(x) &= 11[\alpha - \beta + (\alpha + \beta + 2)x_k]^2 - 8x_k[\alpha - \beta + (\alpha + \beta + 2)x_k] \\
&\quad - 4(\alpha + \beta + 2)(1 - x_k^2) \\
&\quad + (x - x_k)\{(7\alpha + 7\beta + 12)[\alpha - \beta + (\alpha + \beta + 2)x_k] \\
&\qquad\qquad - 4x_k(\alpha + \beta + 2)\} \\
&\quad - \frac{2(x - x_k)}{1 - x_k^2}[\alpha - \beta + (\alpha + \beta + 2)x_k]\{3[\alpha - \beta + (\alpha + \beta + 2)x_k]^2 \\
&\qquad\qquad - 7x_k(\alpha - \beta + (\alpha + \beta + 2)x_k] + 4x_k^2\}.
\end{aligned}
\right.
$$

Let $|x| \leqq 1 - \eta$, $|x - x_k| \leqq \delta \leqq \frac{1}{4}\eta$ $(0 < \eta, \delta < 1)$; then, for $k = 1, 2, \cdots, n$,

$$
\frac{|x - x_k|}{1 - x_k^2} \leqq \frac{\delta}{1 - x^2 + x^2 - x_k^2} \leqq \frac{\delta}{2\eta - \eta^2 - 2\delta} \leqq \frac{\delta}{\eta - 2\delta} \leqq \frac{\delta}{2\delta} = \frac{1}{2}.
$$

Thus, for $|x| \leqq 1-\eta$, $|x-x_k| \leqq \delta \leqq \frac{1}{4}\eta$ $(0<\delta,\ \eta<1)$ and $k = 1, 2, \cdots, n$,

$$(18)\quad \begin{cases} \mu_1(x) \geqq 2\eta - \eta^2 - \delta[2|\alpha-\beta| + |2\alpha+2\beta+3|] \geqq \eta - \nu_1\delta, \\ \qquad\qquad\qquad \nu_1 \equiv 2|\alpha-\beta| + |2\alpha+2\beta+3|, \\ \mu_2(x) \geqq -8|\alpha-\beta| - 12(\alpha+\beta+2) \\ \qquad - (|\alpha-\beta|+\alpha+\beta+2)[3(|\alpha-\beta|+\alpha+\beta+2)^2 \\ \qquad + 7(|\alpha-\beta|+\alpha+\beta+2)+4] \\ \qquad - \delta[|7\alpha+7\beta+12|(|\alpha-\beta|+\alpha+\beta+2) \\ \qquad + 4(\alpha+\beta+2)] \equiv -\nu_2, \\ t(x) \geqq 4n(n+\alpha+\beta+1)(\eta-\nu_1\delta) - \nu_2 \geqq n+\alpha+\beta+1-\nu_2 > 0 \\ \left(|x| \leqq 1-\eta,\ |x-x_k| \leqq \delta \leqq \min\left\{\frac{1}{4}\eta, \frac{\eta}{2\nu_1}\right\},\ n \geqq \max\left\{\frac{1}{2\eta}, \nu_2+1\right\}\right), \\ s(x) \geqq 1 - x^2 + x^2 - x_k^2 - 2|x-x_k|(|\alpha-\beta|+\alpha+\beta+2) \\ \qquad\qquad\qquad \geqq \eta - \nu_3\delta \geqq 0, \\ \qquad\qquad \nu_3 \equiv 2(|\alpha-\beta|+\alpha+\beta+3)\left(\delta \leqq \frac{\eta}{\nu_3}\right). \end{cases}$$

We have thus established

LEMMA 1. *For general Jacobi abscissas, $u_{k,n}(x) \geqq 0$ $(k = 1, 2, \cdots, n)$ for $|x| \leqq 1-\eta$, $|x-x_k| \leqq \delta \leqq \min\{\frac{1}{4}\eta, \eta/2\nu_1, \eta/\nu_3\}$, if n is sufficiently large: $n \geqq \max\{1/2\eta, \nu_2+1\}$. Here, ν_1, ν_2, ν_3 depend on α and β only* (see (18)).

We also need

LEMMA 2. *For general Jacobi abscissas, given η and δ $(0<\eta, \delta<1)$, we have uniformly in $[-1+\eta, 1-\eta]$*

$$\lim_{n\to\infty} \sum_{|x-x_{k,n}|>\delta} |u_{k,n}(x)|\, l_{k,n}^4(x) = 0, \qquad \lim_{n\to\infty} \sum_{|x-x_{k,n}|\leqq\delta} u_{k,n}(x) l_{k,n}^4(x) = 1.$$

Proof. By (17), we have uniformly in $[-1, 1]$

$$(19)\qquad |u_k(x)| \leqq |p(x)| + \frac{(x-x_k)^2|t(x)|}{6(1-x_k^2)^2} = \frac{O(1)}{(1-x_k^2)^3} + \frac{O(n^2)}{(1-x_k^2)^2}.$$

Furthermore, for $\omega(x) = P_n^{(\alpha,\beta)}(x)$, we have ([8], pp. 164, 232)

$$(20)\quad \begin{cases} P_n^{(\alpha,\beta)}(x) = O(n^{-\frac{1}{2}}), & \text{for } |x| \leqq 1-\eta, \\ \left.\begin{array}{l} |\omega'(x_k)| > \tau(n-k+1)^{-\alpha-\frac{3}{2}} n^{\alpha+2},\ \tau k^{-\beta-\frac{3}{2}} n^{\beta+2} \\ 1 - x_k^2 > \dfrac{\tau(n-k+1)^2}{n^2}, \dfrac{\tau k^2}{n^2} \end{array}\right\}, & \begin{array}{r} \text{according as } 0 \leqq x_k < 1 \\ \text{or } -1 < x_k \leqq 0, \end{array} \end{cases}$$

where τ, different in different formulas, generally denotes a constant depending only on α and β, but independent of x, n (here, also independent of k).[4] Thus (19), combined with (2), gives uniformly in $[-1 + \eta, 1 - \eta]$

$$| u_k(x) | = O\left(\frac{n^6}{(n - k + 1)^4}\right), O\left(\frac{n^6}{k^4}\right),$$

according as $0 \leqq x_k < 1$ or $-1 < x_k \leqq 0$,

$$S \equiv \sum_{|x-x_k|>\delta} | u_k(x) | l_k^4(x) < O(1) \sum_{k=1}^{n} \frac{n^4}{k^4} \frac{k^{4\alpha+6}}{n^{4\alpha+8}}$$

$$+ O(1) \sum_{k=1}^{n} \frac{n^4}{(n - k + 1)^4} \frac{(n - k + 1)^{4\beta+6}}{n^{4\beta+8}} \equiv \sum{}' + \sum{}'',$$

$$\sum{}' = O\left(\frac{1}{n}\right), O\left(\frac{\log n}{n}\right), O\left(\frac{1}{n^{4+4\alpha}}\right),$$

according as $\alpha \geqq -\frac{1}{2}$, $-\frac{3}{4} \leqq \alpha < -\frac{1}{2}$, $-1 < \alpha < -\frac{3}{4}$,

and similarly for $\sum''$. The first statement of our lemma now follows; we make use of (9) for its second statement. In view of Theorem A, we have

THEOREM II. *Let $f(x)$ be an arbitrary continuous function on* $[-1, 1]$. *For general Jacobi abscissas,* $\lim_{n\to\infty} F_n[f] = f(x)$ *uniformly in* $-1 + \eta \leqq x \leqq 1 - \eta$, *that is, in any fixed interval wholly inside* $(-1, 1)$.

However, for trigonometric abscissas ($\alpha = \beta = -\frac{1}{2}$) (Theorem I), the above convergence is uniform on the *entire* interval $[-1, 1]$. We therefore seek to extend Theorem I to other Jacobi abscissas. Our considerations are confined, in view of Theorem II, to $x \in [1 - \eta, 1]$ and $[-1, -1 + \eta]$, $\eta > 0$ fixed, sufficiently small, and we shall derive results for $x \in [1 - \eta, 1]$ only, since interchanging α and β reduces the discussion of the second interval to that of the first.

LEMMA 3. *For Jacobi abscissas, with* $-\frac{1}{4} > \alpha \geqq -\frac{1}{2}$ *and* $0 > \beta > -1$, *where* $\alpha - \beta < \frac{1}{2}$ *if* $-\frac{3}{4} > \beta > -1$, *we have uniformly on* $[1 - \eta, 1]$

$$\lim_{n\to\infty} \sum_{|x-x_{k,n}|>\delta} | u_{k,n}(x) | l_{k,n}^4(x) = 0, \qquad (0 < \delta < \eta < \tfrac{1}{4}).$$
$$\sum_{|x-x_{k,n}|\leqq\delta} | u_{k,n}(x) | l_{k,n}^4(x) = O(1)$$

Proof. Here and hereafter, $x \in [1 - \eta, 1]$. From (17), for $x_k \geqq x$ and $| x - x_k | \leqq \delta < \frac{1}{3} < \dfrac{1}{2\beta + 3}$,

$$p(x) \geqq 1 + \frac{2(x - x_k)(\beta + 1)}{1 + x_k} \geqq 1 + 2(x - x_k)(\beta + 1) \geqq 1 - 2(\beta + 1)\delta > \delta;$$

[4] In (20), our x_{n-k+1} is Szegö's x_k.

while, for $x > x_k$, $p(x) > -\frac{1}{2}$, since $p(x)$ is linear and

$$p(x_k) = 1, \qquad p(1) = 1 - 2(\alpha + 1) + 2(\beta + 1)\,\frac{1 - x_k}{1 + x_k} > -2\alpha - 1 > -\tfrac{1}{2}.$$

Thus, $p(x) > -\frac{1}{2}$ for $|x - x_k| \leqq \delta$; $|p(x)| < p(x) + 1$. In the same way, for $|x - x_k| \leqq \delta < \eta < \frac{1}{4}$,

$$\mu_1(x) \geqq 1 - x_k^2 + (x_k - x)[-(1 + x_k) + 3x_k] > 1 - x_k^2 > 0 \quad \text{for } x_k > x;$$

$\mu_1(x) > 0$ for $x > x_k$, since $\mu_1(x_k) = 1$, $\mu_1(x)$ is linear and

$$\mu_1(1) \geqq \begin{cases} (1 - x_k)[-1 + \frac{1}{2}(1 + x_k) + \frac{1}{2}(1 - x_k)] = 0 \\ \qquad\qquad\qquad\qquad \text{for} \quad \beta \geqq -\frac{3}{4}, \\ (1 - x_k)[-1 - 2\alpha(1 + x_k) + 2(\alpha + \frac{1}{2})(1 - x_k)] > 0 \\ \qquad\qquad\qquad\qquad \text{for} \quad -\frac{3}{4} > \beta > -1,\ \alpha - \beta < \frac{1}{2}. \end{cases}$$

Thus, $\mu_1(x) \geqq 0$ for $|x - x_k| \leqq \delta < \frac{1}{4}$, with α and β as in the lemma.

Now, from (17) and (20), letting $\mu(x) = 4n(n + \alpha + \beta + 1)\mu_1(x)$, we get

$$\begin{aligned} |u_k(x)| &\leqq p(x) + 1 + \frac{1}{6}\frac{(x - x_k)^2}{(1 - x_k^2)^2}\{t(x) - [t(x) - \mu(x)]\} \\ &\qquad\qquad + \frac{1}{6}\frac{(x - x_k)^2}{(1 - x_k^2)^2}\,|t(x) - \mu(x)| \\ &\leqq u_k(x) + 1 + \frac{1}{3}\frac{(x - x_k)^2}{(1 - x_k^2)^2}\,|t(x) - \mu(x)| \\ &= u_k(x) + 1 + O(1)\frac{(x - x_k)^2}{(1 - x_k^2)^2} + O(1)\frac{|x - x_k|^3}{(1 - x_k^2)^3} \\ &\qquad\qquad\qquad\qquad\qquad\qquad (|x - x_k| \leqq \delta). \end{aligned} \tag{21}$$

By (19), and since we have ([8], p. 164) uniformly in $[1 - \eta, 1]$,

$$|\omega(x)| = |P_n^{(\alpha,\beta)}(x)| = O(n^a), \qquad a = \max\{\alpha, -\tfrac{1}{2}\}, \tag{22}$$

$$\begin{aligned} \sum\nolimits_1 \equiv \sum_{|x - x_k| > \delta} |u_k(x)|\, l_k^4(x) &= O(1) \sum_{|x - x_k| > \delta} |u_k(x)| \left|\frac{\omega(x)}{\omega'(x_k)}\right|^4 \\ &= O(1)\sum_{k=1}^{n} \frac{n^6}{k^4}\, n^{4a}\left(\frac{k^{4\alpha+6}}{n^{4\alpha+8}} + \frac{k^{4\beta+6}}{n^{4\beta+8}}\right) \\ &\equiv \sum{}' + \sum{}'', \end{aligned}$$

$$\sum{}' = \frac{O(1)}{n^2}\sum_{k=1}^{n} k^{4\alpha+2} = O(n^{4\alpha+1}); \tag{23}$$

$$\sum{}'' = \begin{cases} O(1)n^{4a}\sum\limits_{k=1}^{n}\left(\dfrac{k}{n}\right)^{4\beta+2} = O(n^{4a+1}), & \text{if } \beta \geqq -\frac{1}{2}, \\ O(1)n^{4a+1}\sum\limits_{k=1}^{n}\left(\dfrac{k}{n}\right)^{4\beta+3}\dfrac{1}{k} = O(n^{4a+1}\log n), & \text{if } -\frac{1}{2} > \beta \geqq -\frac{3}{4}, \\ O(1)n^{4a-4\beta-2}\sum\limits_{k=1}^{n} k^{4\beta+2} = O(n^{4a-4\beta-2}), & \text{if } -\frac{3}{4} > \beta > -1. \end{cases}$$

Hence, $\sum_1 = o(1)$ under the restrictions of the lemma on α and β, and its first statement is established.

Turning to the second statement of the lemma, we see first from (9) and (23) that uniformly in $[1 - \eta, 1]$

$$\sum_{|x-x_k| \leqq \delta} u_k(x) l_k^4(x) = 1 - \sum_{|x-x_k| > \delta} u_k(x) l_k^4(x) = 1 + o(1).$$

Thus, using (21), (20) and (22), we get

$$\sum\nolimits_2 \equiv \sum_{|x-x_k| \leqq \delta} | u_k(x) | l_k^4(x) < 1 + o(1) + \left(\sum_{k=1}^{n} l_k^2(x)\right)^2$$
$$+ O(1) \sum_{|x-x_k| \leqq \delta} \frac{n^4}{k^4} n^{2\alpha} \frac{k^{2\alpha+3}}{n^{2\alpha+4}} l_k^2(x) + O(1) \sum_{|x-x_k| \leqq \delta} \frac{n^6}{k^6} \frac{k^{3\alpha+\frac{9}{2}}}{n^{3\alpha+6}} n^{3\alpha} | l_k(x) |.$$

Observe further that for $0 > \alpha, \beta > -1$ the Jacobi abscissas are "strongly normal" ([2], p. 12), so that by a theorem of Fejér [2],

$$\sum_{k=1}^{n} l_k^2(x) \leqq \max\left\{-\frac{1}{\alpha}, -\frac{1}{\beta}\right\} \equiv \mu \qquad \text{uniformly in } [-1, 1].$$

Therefore, uniformly in $[1 - \eta, 1]$

$$\sum\nolimits_2 < 1 + o(1) + \mu^2 + O(1) \sum_{k=1}^{n} k^{2\alpha-1} l_k^2(x) + O(1) \sum_{k=1}^{n} k^{3\alpha-\frac{3}{2}} | l_k(x) |$$
$$< O(1) + O(1) \left(\sum_{k=1}^{n} k^{6\alpha-3} \sum_{k=1}^{n} l_k^2(x)\right)^{\frac{1}{2}} = O(1).$$

This completes the proof of Lemma 3.

Lemma 4. *For general Jacobi abscissas*

$$4(1 - x_{k,n}^2) n(n + \alpha + \beta + 1) \geqq [\alpha - \beta + (\alpha + \beta + 2) x_{k,n}]^2$$
$$+ 8 x_{k,n} [\alpha - \beta + (\alpha + \beta + 2) x_{k,n}] + 4(\alpha + \beta + 2)(1 - x_{k,n}^2)$$
$$(k = 1, 2, \cdots, n).$$

Proof. A theorem of Laguerre [4] states that if $\phi(x)$ is a polynomial of degree n, whose roots are all real and distinct, then for any such root x_0

$$3(n - 2)[\phi''(x_0)]^2 - 4(n - 1)\phi'(x_0)\phi'''(x_0) \geqq 0.$$

A direct application to the Jacobi polynomial, if (16) and its derived equation are used, and some elementary transformations yield the desired inequality.

Lemma 5. *For Jacobi abscissas, with* $-\frac{3}{4} \leqq \alpha \leqq -\frac{1}{2}$ *and* $-1 < \beta \leqq -\frac{1}{4}$, $u_{k,n}(x) \geqq 0$ *in* $[1 - \eta, 1]$ *for* $x_{k,n}$ *such that* $| x - x_{k,n} | \leqq \delta$ $(0 < \delta < \eta$; η, δ *fixed, sufficiently small, as specified below).*

Proof. From (17), for $|x - x_k| \leqq \delta < \eta < \frac{1}{2}$, $p(x) \geqq 0$ since

$$(24) \quad p(x) \geqq \begin{cases} 1 - \dfrac{2(\alpha+1)}{1-x_k}(x - x_k) \geqq 1 - 2(\alpha+1) \geqq 0, & \text{if } x_k \leqq x, \\ 1 - \dfrac{2(\beta+1)}{1+x_k}(x_k - x) > 1 - 2(\beta+1)\delta > 0, & \text{if } x_k > x. \end{cases}$$

Also, $\mu_1(x) \geqq x_k(1 - x_k)$ for $x_k \leqq x$, since

$$\mu_1(x_k) = 1 - x_k^2 > x_k(1 - x_k), \quad \mu_1(1) \geqq (1 - x_k)[-1 + (1 + x_k)] = x_k(1 - x_k).$$

For $x_k > x$, we have

$$\mu_1(x) \geqq 1 - x_k^2 + (x_k - x)[-\tfrac{3}{2}(1 + x_k) + \tfrac{1}{2}(1 - x_k) + 3x_k] \geqq (1 - x_k)(1 + x).$$

Thus, for $|x - x_k| \leqq \delta$,

$$(25) \quad \mu_1(x) \geqq \begin{cases} x_k(1 - x_k) = (1 - x_k^2)\dfrac{x_k}{1 + x_k} \geqq (\tfrac{1}{2} - \eta)(1 - x_k^2), & \text{if } x_k \leqq x, \\ (1 - x_k^2)\dfrac{1 + x}{1 + x_k} \geqq (1 - x_k^2)(1 - \delta), & \text{if } x_k > x. \end{cases}$$

In view of (24), it remains only to establish that $t(x) \geqq 0$ for $|x - x_k| \leqq \delta$, $x \neq x_k$. This we do by means of (25) and Lemma 4. Here we consider the cases $x_k < x$, $x_k > x$ separately.

1°. $x_k > x$: $\mu_1(x) \geqq (1 - x_k^2)(1 - \delta)$. From Lemma 4,

$$\begin{aligned} 4n(n + \alpha + \beta + 1)\mu_1(x) &\geqq (1 - \delta)\{4(\alpha + \beta + 2)(1 - x_k^2) \\ &\quad + 8x_k[\alpha - \beta + (\alpha + \beta + 2)x_k] + [\alpha - \beta + (\alpha + \beta + 2)x_k]^2\} \\ &\geqq 4(\alpha + \beta + 2)(1 - x_k^2) + 8x_k[\alpha - \beta + (\alpha + \beta + 2)x_k] \\ &\quad + [\alpha - \beta + (\alpha + \beta + 2)x_k]^2 - 17\delta. \end{aligned}$$

Hence,

$$\begin{aligned} t(x) \geqq\ & 12(\alpha + 1)^2(1 + x_k)^2 \\ & + \frac{2(x_k - x)(\alpha + 1)}{1 - x_k}[3(\alpha + 1)^2(1 + x_k)^2 - 7(\alpha + 1)(1 + x_k) + 4] \\ & - 18\eta - 45\delta. \end{aligned}$$

Now, $0 < \frac{1}{2} - \frac{1}{2}\eta < \frac{1}{4}(1 + x_k) \leqq (\alpha + 1)(1 + x_k) \equiv m \leqq \frac{1}{2}(1 + x_k) < 1$ and $3m^2 - 7m + 4 \geqq 0$ over $[\frac{1}{2} - \frac{1}{2}\eta, 1]$. Thus,

$$t(x) \geqq \tfrac{3}{4}(1 + x_k)^2 - 18\eta - 45\delta \geqq 3 - 24\eta - 45\delta > 3 - 69\eta > 0$$
$$(0 < \delta < \eta < \tfrac{1}{23}).$$

2°. $x_k < x$: $\mu_1(x) \geqq (\frac{1}{2} - \eta)(1 - x_k^2)$. Proceeding as in the first case, we have

$$4n(n + \alpha + \beta + 1)\mu_1(x) \geqq \tfrac{1}{2}\{4(\alpha + \beta + 2)(1 - x_k^2) \\ + 8x_k[\alpha - \beta + (\alpha + \beta + 2)x_k] + [\alpha - \beta + (\alpha + \beta + 2)x_k]^2\} - 17\eta,$$

$$t(x) \geqq \tfrac{23}{2}(\alpha + 1)^2(1 + x_k)^2 - 4x_k(\alpha + 1)(1 + x_k) \\ + \frac{2(x - x_k)(\alpha + 1)}{1 - x_k}[-3(\alpha + 1)^2(1 + x_k)^2 + 7(\alpha + 1)(1 + x_k) - 4] \\ - 43\eta - 45\delta.$$

Now, $-3m^2 + 7m - 4$ is minimum over $[\frac{1}{4}(1 + x_k), 1]$ at $m = \frac{1}{4}(1 + x_k)$, so that

$$t(x) \geqq \tfrac{23}{2}(\alpha + 1)^2(1 + x_k)^2 - \tfrac{1}{2}x_k(\alpha + 1)(1 + x_k) \\ - \tfrac{3}{8}(\alpha + 1)(1 + x_k)^2 - 8(\alpha + 1) - 88\eta \\ \geqq \tfrac{23}{8} - \tfrac{21}{8} - 111\eta = \tfrac{1}{4} - 111\eta > 0 \qquad (0 < \delta < \eta < \tfrac{1}{444}).$$

LEMMA 6. *For Jacobi abscissas, with* $-\frac{3}{4} \leqq \alpha \leqq -\frac{1}{2}$ and $-1 < \beta \leqq -\frac{1}{4}$, *we have uniformly in* $[1 - \eta, 1]$

$$\sum_{|x - x_{k,n}| > \delta} |u_{k,n}(x)| l_{k,n}^4(x) = o(1), \qquad \sum_{|x - x_{k,n}| \leqq \delta} |u_{k,n}(x)| l_{k,n}^4(x) = 1 + o(1) \\ (0 < \delta < \eta < \tfrac{1}{444}).$$

Proof. In view of (9) and Lemma 5, we need only prove the first statement. Again using (19), (22) and (20), we get

$$\sum \equiv \sum_{|x - x_k| > \delta} |u_k(x)| l_k^4(x) = O(1) \sum_{|x - x_k| > \delta} \frac{n^6}{k^4} \left|\frac{\omega(x)}{\omega'(x_k)}\right|^4 \\ = O(1) \frac{1}{n^2} \sum_{k=1}^{n} \frac{n^6}{k^4} \left(\frac{k^{4\alpha+6}}{n^{4\alpha+8}} + \frac{k^{4\beta+6}}{n^{4\beta+8}}\right) \equiv \sum{}' + \sum{}''.$$

We readily obtain

$$\sum{}' = O\left(\frac{1}{n}\right) \quad (\alpha = -\tfrac{1}{2}), \qquad O\left(\frac{\log n}{n}\right) \quad (-\tfrac{3}{4} \leqq \alpha < -\tfrac{1}{2});$$

$$\sum{}'' = O\left(\frac{1}{n}\right) \quad (-\tfrac{1}{4} \geqq \beta \geqq -\tfrac{1}{2}), \qquad O\left(\frac{\log n}{n}\right) \quad (-\tfrac{3}{4} \leqq \beta < -\tfrac{1}{2}), \\ O(n^{-4\beta-4}) \quad (-1 < \beta < -\tfrac{3}{4}).$$

Thus, $\sum = o(1)$, and our lemma is established.

Lemmas 3 and 6 and Theorem A permit us to supplement Theorem II

THEOREM III. *Let $f(x)$ be an arbitrary continuous function on* $[-1, 1]$. *For Jacobi abscissas,* $\lim_{n \to \infty} F_n[f] = f(x)$ *uniformly:*

(i) *on* $[-1 + \eta, 1]$ if $-\frac{1}{4} > \alpha \geqq -\frac{1}{2}$ *and* $0 > \beta > -1$, *provided that* $\alpha - \beta < \frac{1}{2}$ *if* $-\frac{3}{4} > \beta > -1$;

(ii) *on* $[-1+\eta, 1]$ *if* $-\frac{1}{2} \geqq \alpha \geqq -\frac{3}{4}$ *and* $-\frac{1}{4} \geqq \beta > -1$;
(iii) *on* $[-1, 1-\eta]$ *if* α *and* β *are interchanged in* (i) *and* (ii);
(iv) *on* $[-1, 1]$ *if* $-\frac{1}{4} > \{\alpha, \beta\} \geqq -\frac{1}{2}$ *or if* $-\frac{1}{2} \geqq \{\alpha, \beta\} \geqq -\frac{3}{4}$.

COROLLARY 3.1. *For symmetric Jacobi abscissas, with* $-\frac{1}{4} > \alpha = \beta \geqq -\frac{3}{4}$, $F_n[f] \to f(x)$ *uniformly on* $[-1, 1]$ *as* $n \to \infty$.

That Theorem III is the "best possible" is illustrated by the following example. Let $\alpha = \beta = -\frac{1}{4}$ and $x = 1$. We can show that there exists a function $f(x)$, continuous on $[-1, 1]$, and such that $F_n(1)$ does not converge to $f(1)$, as $n \to \infty$. For, (2) and (17) yield, when $\alpha = \beta = -\frac{1}{4}$,

$$(26) \quad u_{k,n}(1) l_{k,n}^4(1) = \left[\frac{n(2n+1)}{3} - \frac{1}{8} + \frac{1}{8(1-x_k)(1-x_k^2)}\right]\left[\frac{\omega(1)}{\omega'(x_k)}\right]^4 \frac{1}{(1-x_k^2)^2}$$

$$(k = 1, 2, \cdots, n;\ n = 1, 2, \cdots).$$

Furthermore, for $\omega(x) = P_n^{(-\frac{1}{4}, -\frac{1}{4})}(x)$, we have ([8], pp. 57, 232)

$$(27) \quad \begin{cases} \omega(1) = \dbinom{n-\frac{1}{4}}{n} > (\tau_1 n)^{-\frac{1}{4}}, \\ \left.\begin{array}{l} |\omega'(x_k)| < \tau_2^{\frac{1}{4}} k^{-\frac{5}{4}} n^{\frac{7}{4}} \\ 1 - x_k^2 < \tau_3^{\frac{1}{2}} k^2 n^{-2} \end{array}\right\} \quad \text{for } 0 \geqq x_k > -1; \text{ i.e., } k = 1, 2, \cdots, [\tfrac{1}{2}(n+1)], \end{cases}$$

where τ_1, τ_2 and τ_3 are absolute constants bounded from zero and infinity, independent of n and k, and determined solely from the fact that $\alpha = \beta = -\frac{1}{4}$. Now, let $f(x)$ be any continuous, monotone non-decreasing function on $[-1, 1]$ with $f(1) = 12\tau_1\tau_2\tau_3$ and $f(0) = 0$. Given $\epsilon > 0$, let $\delta > 0$ be such that $|f(1) - f(x)| < \epsilon$ for $|1 - x| < \delta$ $(-1 \leqq x \leqq 1)$. Then, from (26) and (27),

$$|f(1) - F_n(1)| = \sum_{k=1}^{n} [f(1) - f(x_k)] u_k(1) l_k^4(1) \geqq \sum_{x_k \leqq 0} {}_k [f(1) - f(x_k)] u_k(1) l_k^4(1)$$

$$\geqq 12\tau_1\tau_2\tau_3 \sum_{k=1}^{[\frac{1}{2}(n+1)]} \frac{2n^2}{3}\left[\frac{\omega(1)}{\omega'(x_k)}\right]^4 \frac{1}{(1-x_k^2)^2}$$

$$> \frac{8}{n^2} \sum_{k=1}^{[\frac{1}{2}(n+1)]} k = \frac{4}{n^2}\left[\frac{n+1}{2}\right]\left(1 + \left[\frac{n+1}{2}\right]\right) > 1.$$

This completes our demonstration.

5. General Laguerre abscissas ($\alpha > -1$). As previously, our starting point is the differential equation for $\omega(x) = L_n^{(\alpha)}(x)$, the orthonormal Laguerre polynomial,

$$(28) \qquad x\omega''(x) + (\alpha + 1 - x)\omega'(x) + n\omega(x) = 0,$$

and its (twice) derived equations, as in §3, (7) being employed additionally to obtain

$$
(29)\quad \begin{cases}
u_k(x) = p(x) + \dfrac{(x - x_k)^2}{6x_k^2}\, t(x), \\[2mm]
p(x) \equiv 1 - \dfrac{2(x - x_k)(x_k - \alpha - 1)}{x_k} \equiv \dfrac{s(x)}{x_k}, \\[2mm]
t(x) = 2(n - 1)\mu_1(x) + \mu_2(x), \\
\mu_1(x) \equiv 2x_k + (-4x_k + 4\alpha + 3)(x - x_k), \\
\mu_2(x) \equiv (x_k - \alpha - 1)\Big\{11(x_k - \alpha - 1) + 4 - (x - x_k) \\
\qquad \times \left[\dfrac{6(x_k - \alpha - 1)^2 + 7(x_k - \alpha - 1) + 2}{x_k} + x_k - \alpha - 1\right]\Big\}.
\end{cases}
$$

We observe that in an arbitrarily fixed positive interval $(0 <)\ h \leqq x \leqq A$,

$$
s(x) \geqq \begin{cases}
x - (x - x_k) - 2x(x - x_k) + 2(\alpha + 1)(x - x_k) \geqq x - \delta - 2x\delta \\
\qquad \geqq \tfrac{1}{2}x - \delta \geqq 0 \quad (x_k \leqq x, |x - x_k| \leqq \delta \leqq \min\{\tfrac{1}{4}, \tfrac{1}{2}h\}), \\
x_k - 2(\alpha + 1)(x_k - x) \geqq h - \delta - 2(\alpha + 1)\delta \geqq 0 \\
\qquad \left(x_k > x, |x - x_k| \leqq \delta \leqq \dfrac{h}{2\alpha + 4} < \tfrac{1}{2}h\right),
\end{cases}
$$

$$
p(x) \geqq 0 \quad \text{in} \quad [h, A] \qquad \left(|x - x_k| \leqq \delta \leqq \min\left\{\frac{1}{4}, \frac{h}{2\alpha + 4}\right\}\right).
$$

As for $t(x)$, we have in $[h, A]$, for $|x - x_k| \leqq \delta$,

$$
(30)\quad \begin{cases}
\mu_1(x) \geqq 2h - \delta\,|-4x_k + 4\alpha + 3| > 2h - \tfrac{1}{2}\nu_1\delta, \\
\qquad \nu_1 \equiv 8(A + h) + 8|\alpha| + 6, \\
\mu_2(x) \geqq -4(\alpha + 1) - \delta(A + \alpha + 1) \\
\qquad \times \left[\dfrac{6(A + \alpha + 1)^2 + 7(A + \alpha + 1) + 2}{h - \delta} + A + \alpha + 1\right] \equiv -\nu_2, \\
t(x) \geqq (n - 1)(4h - \nu_1\delta) - \nu_2 \geqq (n - 1)\,\dfrac{4h}{\nu_1 + 1} - \nu_2 \geqq 0 \\
\qquad \left(h \leqq x \leqq A, |x - x_k| \leqq \delta \leqq \dfrac{4h}{\nu_1 + 1}, n \geqq 1 + \dfrac{\nu_2(\nu_1 + 1)}{4h}\right).
\end{cases}
$$

Thus, we may state

LEMMA 7. *For general Laguerre abscissas, $u_{k,n}(x) \geqq 0$ in $[h, A]$, provided*

$$
|x - x_{k,n}| \leqq \delta \leqq \min\left\{\frac{1}{4}, \frac{h}{2\alpha + 4}, \frac{4h}{\nu_1 + 1}\right\}, \qquad n \geqq 1 + \frac{\nu_2(\nu_1 + 1)}{4h}.
$$

Here ν_1, ν_2 are independent of k, n and x (see (30)).

LEMMA 8. *For general Laguerre abscissas, we have uniformly in* $[h, A]$, *with* $0 < \delta < h$ $(n \to \infty)$

$$\sum_{|x-x_{k,n}|>\delta} |u_{k,n}(x)| \, l_{k,n}^4(x) = O(n^a) = o(1), \qquad a = \max\{-\alpha - 1, -\tfrac{1}{2}\};$$

$$\sum_{|x-x_{k,n}|\leqq\delta} u_{k,n}(x) l_{k,n}^4(x) = 1 + o(1).$$

Proof. The method applied below was first introduced by Shohat [5] and also used by Szegö ([8], p. 337). We observe that in (29), for $h \leqq x \leqq A$,

(31) $u_k(x) = O\left(\dfrac{1}{x_k^3}\right) + O\left(\dfrac{n}{x_k^2}\right)$ if x_k is "small" $(x_k < h - \delta)$;

(32) $u_k(x) = O(x_k^{\frac{3}{2}}) + O(nx_k^{\frac{1}{2}})$ if x_k is "large" $(x_k > A + \delta)$;

(33) $u_k(x) = O(n)$ for "intermediate" x_k $(h - \delta \leqq x_k \leqq A + \delta)$.

1°. Let x_k be small: $x_k < h - \delta$. Then, necessarily $|x - x_k| > \delta$, and

$$(34) \qquad \sum{}' \equiv \sum_{x_k<h-\delta}{}_k |u_k(x)| \, l_k^4(x) = O(n^{-1}) \sum_{x_k<h-\delta}{}_k \left(\frac{1}{x_k^3} + \frac{n}{x_k^2}\right) \frac{1}{[\omega'(x_k)]^4},$$

since ([5], p. 132)

$$(35) \qquad |\omega(x)| = O(n^{-\frac{1}{4}}) \qquad \text{uniformly in } [h, A].$$

Observe, further, that ([8], p. 233)

$$(36) \qquad \frac{1}{x_{k,n}} = O(n) \qquad (0 < x_k \leqq A).$$

We now introduce $\lambda_{k,n}$ $(\equiv H_{k,n}$; see [5, 6]) $(k = 1, 2, \cdots, n)$, the coefficients of the Gaussian mechanical quadrature formula corresponding to the Laguerre abscissas ([5], p. 134):

$$(37) \qquad \begin{aligned} \lambda_{k,n} &= \int_0^\infty l_{k,n}(x) e^{-x} x^\alpha \, dx = \frac{1}{x_{k,n}[\omega'(x_{k,n})]^2}; \\ \sum_{k=1}^n x_{k,n}^m \lambda_{k,n} &= \int_0^\infty x^m e^{-x} x^\alpha \, dx \qquad (m \text{ integer} \leqq 2n - 1). \end{aligned}$$

Employing ([6], p. 210) (where α corresponds to our $\alpha + 1$) we obtain

$$(38) \quad \lambda_{k,n} < \tau n^{-\alpha-1} \; (-1 < \alpha \leqq -\tfrac{1}{2}), \qquad \tau n^{-\frac{1}{2}} \; (\alpha \geqq -\tfrac{1}{2}) \; (k = 1, 2, \cdots, n),$$

$$\sum{}' = O(n^{-1}) \sum{}' n\lambda_{k,n}^2 = O(1) \sum{}' \lambda_{k,n}^2 < O(1) \sum_{k=1}^n \lambda_{k,n}^2$$

$$= O(n^a) \sum_{k=1}^n \lambda_{k,n} = O(n^a) \int_0^\infty e^{-x} x^\alpha \, dx = O(n^a), \qquad a = \max\{-\alpha - 1, -\tfrac{1}{2}\}.$$

2°. Let x_k be large: $x_k > A + \delta$. Then, necessarily $|x - x_k| > \delta$, and

$$\sum{}''' \equiv \sum_{x_k > A+\delta} |u_k(x)| l_k^4(x) = O(1) \sum{}''' (x_k^3 + nx_k^2) l_k^4(x)$$

(39)

$$< O(1) \sum{}''' \frac{nx_k^2}{(x - x_k)^4} \left[\frac{\omega(x)}{\omega'(x_k)}\right]^4$$

since ([8], p. 125)

(40) $$x_{k,n} < \tau k \quad \text{for large } k.$$

We remark that

$$\frac{x_k^2}{(x - x_k)^4} = \frac{1}{x_k^2\left(1 - \frac{x}{x_k}\right)^4} < \frac{1}{x_k^2\left(1 - \frac{A}{A + \delta}\right)^4} = O(1)\frac{1}{x_k^2},$$

so that, by using (35) again, we get

$$\sum{}''' < O(1) \sum_{k=1}^{n} \lambda_{k,n}^2 = O(n^a) \sum_{k=1}^{n} \lambda_{k,n} = O(n^a), \qquad a = \max\{-\alpha - 1, -\tfrac{1}{2}\}.$$

3°. Let x_k be intermediate: $h - \delta \leqq x_k \leqq A + \delta$. Then,

$$\sum{}'' \equiv \sum_{|x - x_k| > \delta} |u_k(x)| l_k^4(x) = O(n) \sum{}'' \left[\frac{\omega(x)}{\omega'(x_k)}\right]^4$$

(41) $$= O(n)O(n^{-1}) \sum{}'' x_{k,n}^2 \lambda_{k,n}^2 < O(n^a) \sum_{k=1}^{n} x_{k,n}^2 \lambda_{k,n}$$

$$= O(n^a) \int_0^\infty x^{\alpha+2} e^{-x}\, dx = O(n^a), \qquad a = \max\{-\alpha - 1, -\tfrac{1}{2}\}.$$

This establishes the first statement of Lemma 8, and the second now follows, by virtue of (9) and the first.

THEOREM IV. *Let $f(x)$ be continuous for every finite $x \geqq 0$. Assume further that $f(x) = O(x^m)$ if $x \to +\infty$ (m arbitrarily large, but fixed positive integer). Then, for general Laguerre abscissas, $F_n[f] \to f(x)$ as $n \to \infty$, uniformly over any fixed positive interval $[h, A]$.*

Proof. Here, we must modify (10) and Theorem A. Take $\delta > 0$ such that $|f(x') - f(x'')| < \epsilon$ for $|x' - x''| \leqq \delta$, $0 \leqq x', x'' \leqq A + \delta$, and such that Lemmas 7 and 8 hold. Now, by (9) and Lemma 7, as $n \to \infty$,

$$|F_n[f] - f(x)| \leqq \sum_{|x-x_k| \leqq \delta} |f(x_k) - f(x)| u_k(x) l_k^4(x)$$

$$+ \sum_{|x-x_k| > \delta} |f(x_k)| |u_k(x)| l_k^4(x) + |f(x)| \sum_{|x-x_k|>\delta} |u_k(x)| l_k^4(x)$$

$$\leqq \epsilon O(1) + Mo(1) + \sum_{|x-x_k|>\delta} |f(x_k)| |u_k(x)| l_k^4(x)$$

$$\left(M = \max_{[0, A+\delta]} |f(x)|\right).$$

Proceeding as in Lemma 8, write

$$\sum_{|x-x_k|>\delta} |f(x_k)| \, |u_k(x)| \, l_k^4(x) = \sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3,$$

where the ranges of summation for $\sum_1$, $\sum_2$ and $\sum_3$ are given in the corresponding sums (34), (39) and (41). Thus, just as in those sums, as $n \to \infty$,

$$\sum\nolimits_1 = Mo(1), \qquad \sum\nolimits_2 \leqq M \sum\nolimits'' |u_k(x)| \, l_k^4(x) = Mo(1),$$

$$\sum\nolimits_3 = O(1) \sum\nolimits''' |f(x_{k,n})| \, \lambda_{k,n}^2 = o(1) \sum\nolimits''' |f(x_{k,n})| \, \lambda_{k,n}$$

$$= o(1) \sum\nolimits''' x_{k,n}^m \lambda_{k,n} < o(1) \sum_{k=1}^{n} x_{k,n}^m \lambda_{k,n} < o(1) \int_0^\infty x^{m+\alpha} e^{-x} \, dx = o(1).$$

Hence,

$$|F_n[f] - f(x)| \leqq \epsilon O(1) + Mo(1) + Mo(1), \qquad \text{uniformly in } [h, A],$$

and the result stated follows.

We proceed to extend the result obtained to the interval $[0, A]$, remarking that we need only consider the interval $[0, h]$. For this purpose, we develop some lemmas.

Lemma 9. *For Laguerre abscissas, with* $-\frac{3}{4} \leqq \alpha \leqq -\frac{1}{2}$, *we have uniformly on* $[0, h]$ $(0 < \delta < \frac{1}{2}h < 1/600)$

$$\sum_{|x-x_{k,n}|>\delta} |u_{k,n}(x)| \, l_{k,n}^4(x) = o(1), \qquad \sum_{|x-x_{k,n}|\leqq\delta} |u_{k,n}(x)| \, l_{k,n}^4(x) = 1 + o(1)$$

$$(n \to \infty).$$

Proof. Applying the theorem of Laguerre used in Lemma 4 to $L_n^{(\alpha)}(x)$ and using (28) and its derived equation, we get

$$3(n-2)(x_k - \alpha - 1)^2 - 4(n-1)(x_k - \alpha - 2)(x_k - \alpha - 1) + 4(n-1)^2 x_k \geqq 0,$$

and by elementary transformations

(42) $$4nx_{k,n} \geqq (x_{k,n} - \alpha - 1)^2 + 4(\alpha + 1) \quad (k = 1, 2, \cdots, n; n = 1, 2, \cdots).$$

Here and hereafter, $0 \leqq x \leqq h < 1/300$. Turning to (29), consider $u_{k,n}(x)$ with $-\frac{3}{4} \leqq \alpha \leqq -\frac{1}{2}$, $|x - x_{k,n}| \leqq \delta < \frac{1}{2}h < 1/600$. There are two cases to consider.

1°. $x_k \leqq x$: Here $p(x) \geqq 1 - 2(x - x_k) \geqq 1 - 2\delta > 0$. Also,

$$\mu_1(x) = 2x_k - 4x_k(x - x_k) - (x - x_k) + 4(\alpha + 1)(x - x_k)$$

$$\geqq 2x_k[1 - 2(x - x_k)] \geqq 2x_k(1 - 2\delta).$$

Furthermore, employing (42) and the fact that $6(\alpha + 1)^2 - 7(\alpha + 1) + 2 \geqq 0$, we obtain

$$
\begin{aligned}
t(x) \geqq{} & (1 - 2\delta)[(x_k - \alpha - 1)^2 + 4(\alpha + 1)] + 11x_k^2 - 22(\alpha + 1)x_k \\
& + 18(\alpha + 1)^2 - 6(\alpha + 1) - 6(\alpha + 1)^3 \\
& + \frac{x}{x_k}(\alpha + 1)[6(\alpha + 1)^2 - 7(\alpha + 1) + 2] \\
& + (x - x_k)[18x_k(\alpha + 1) - 18(\alpha + 1)^2 + 7(\alpha + 1) - 6x_k^2] \\
& > 19(\alpha + 1)^2 - 2(\alpha + 1) - 6(\alpha + 1)^3 - 29h \geqq \tfrac{19}{32} - 29h > 0 \\
& \hspace{20em} (|\, x - x_k \,| \leqq \delta).
\end{aligned}
$$

2°. $x_k > x$: Here $p(x) > 0$, since $p(x_k) = 1$ and $p(0) = 1 + 2x_k - 2(\alpha + 1) > 0$ and $p(x)$ is linear. Similarly, $\mu_1(x) \geqq x_k + 4x_k^2$, since $\mu_1(x)$ is also linear and $\mu_1(x_k) = 2x_k \geqq x_k + 4x_k^2 + x_k(1 - 8h)$, $\mu_1(0) = 3x_k + 4x_k^2 - 4(\alpha + 1)x_k \geqq x_k + 4x_k^2$. Again employing (42), we have

$$t(x) > \tfrac{37}{2}(\alpha + 1)^2 - 4(\alpha + 1) - 6(\alpha + 1)^3 - 18h \geqq \tfrac{1}{16} - 18h > 0.$$

We may summarize: For Laguerre abscissas, with $-\frac{3}{4} \leqq \alpha \leqq -\frac{1}{2}$, we have $u_{k,n}(x) \geqq 0$ in $[0, h]$, provided $|\, x - x_{k,n} \,| \leqq \delta < \frac{1}{2}h < 1/600$.

Now, if we use (35) and (36), the former still holding uniformly in $[0, h]$ for $-\frac{3}{4} \leqq \alpha \leqq -\frac{1}{2}$, we proceed as in Lemma 8.

Lemma 10. *For Laguerre abscissas, with* $-\frac{1}{2} < \alpha < -\frac{1}{4}$, *we have uniformly in* $[0, h]$ $(0 < \delta < \frac{1}{2}h < 1/600)$

$$\sum_{|x - x_{k,n}| > \delta} |\, u_{k,n}(x) \,|\, l_{k,n}^4(x) = o(1), \qquad \sum_{|x - x_{k,n}| \leqq \delta} |\, u_{k,n}(x) \,|\, l_{k,n}^4(x) = O(1)$$

$$(n \to \infty).$$

Proof. Here and hereafter, $0 \leqq x \leqq h < 1/300$. From (29), for $|\, x - x_k \,| \leqq \delta < \frac{1}{2}h < \frac{1}{600}$ and $-\frac{1}{2} < \alpha < -\frac{1}{4}$, we have $|\, p(x) \,| < p(x) + 1$; since the linearity of $p(x)$, $p(x_k) = 1$ and $p(0) = 1 + 2x_k - 2(\alpha + 1) > -2\alpha - 1 > -\frac{1}{2}$ imply that $p(x) > -\frac{1}{2}$ for $x_k > x$; while, for $x_k \leqq x$, $p(x) \geqq 1 - 2(x - x_k) \geqq 1 - 2\delta > 0$. Also, for $|\, x - x_k \,| \leqq \delta < \frac{1}{2}h$,

$$
(43) \quad \mu_1(x) \geqq \begin{cases} 2x_k + (1 - 4x_k)(x - x_k) > 2x_k, & \text{if } x_k \leqq x, \\ 2x_k - 4x_k(x - x_k) + 2(x - x_k) > 4x_k^2 + 2x(1 - 3h) & \\ \qquad = 2x_k\left[2x_k + \dfrac{x}{x_k}(1 - 3h)\right], & \text{if } x_k > x. \end{cases}
$$

Moreover, $t(x) > 0$ for $|\, x - x_k \,| \leqq \delta$, in view of (42) and (43). Thus,

$$
(44) \quad \begin{aligned}
|\, u_k(x) \,| &= \left| p(x) + \frac{(x - x_k)^2}{6x_k^2}\, t(x) \right| \\
&< p(x) + \frac{(x - x_k)^2}{6x_k^2}\, t(x) + 1 = u_k(x) + 1 \quad (|\, x - x_k \,| \leqq \delta).
\end{aligned}
$$

Now, as in Lemma 9,

$$\sum \equiv \sum_{|x-x_{k,n}|>\delta} |\, u_{k,n}(x) \,|\, l_{k,n}^4(x) \equiv \sum' + \sum'' + \sum''',$$

where $\sum'$, $\sum''$ and $\sum'''$ are as indicated in (34), (39) and (41). Using ([5], p. 132),

$$(45) \qquad |\, \omega(x) \,| = O(n^a), \qquad a = \max\left\{\frac{\alpha}{2}, -\frac{1}{4}\right\} \qquad (0 \leqq x \leqq h),$$

and (36), (37), (38), we obtain

$$\sum' = O(n^{2a}) \sum' \left(\frac{1}{x_k^3} + \frac{n}{x_k^2}\right) \frac{1}{[\omega'(x_k)]^4} = O(n^{2a+1}) \sum_{k=1}^{n} \lambda_{k,n}^2 = O(n^{2a+\frac{1}{2}}),$$

$$\sum'' = O(n^{2a}) \sum'' n x_k^2 \lambda_{k,n}^2 = O(n^{2a+\frac{1}{2}}) \sum_{k=1}^{n} \lambda_{k,n} = O(n^{2a+\frac{1}{2}}),$$

$$\sum''' = O(n^{2a}) \sum''' \frac{n x_k^2}{(x-x_k)^4} \frac{1}{[\omega'(x_k)]^4} = O(n^{2a+1}) \sum_{k=1}^{n} \lambda_{k,n}^2 = O(n^{2a+\frac{1}{2}}),$$

which establishes the first statement of our lemma. As an immediate consequence, we have uniformly in $[0, h]$ $(0 < \delta < \frac{1}{2}h < 1/600)$

$$(46) \qquad \sum_{|x-x_k|\leqq\delta} u_k(x) l_k^4(x) = 1 - \sum_{|x-x_k|>\delta} u_k(x) l_k^4(x) = 1 + o(1) \quad (n \to \infty).$$

We observe that from the theory of Hermite interpolation ([8], pp. 323, 324) (see (5))

$$\sum_{k=1}^{n} \left[1 - \frac{\omega''(x_{k,n})}{\omega'(x_{k,n})}(x - x_{k,n})\right] l_{k,n}^2(x) \equiv \sum_{k=1}^{n} \rho_{k,n}(x) l_{k,n}^2(x) \equiv 1 \qquad (n = 1, 2, \cdots).$$

For Laguerre abscissas, from (28), for $0 \leqq x \leqq h < 1/300$ and $\alpha < -\frac{1}{4}$,

$$\rho_{k,n}(x) = 1 - \frac{x_k - \alpha - 1}{x_k}(x - x_k)$$

$$= -\alpha + \frac{x(\alpha+1)}{x_k} - (x - x_k) > \tfrac{1}{4} - x > \tfrac{1}{5}.$$

Thus,

$$(47) \qquad \sum_{k=1}^{n} l_{k,n}^2(x) < 5 \qquad (n = 1, 2, \cdots; 0 \leqq x \leqq h < \tfrac{1}{300}, \alpha < -\tfrac{1}{4}).$$

Now, from (37), (44), (45), (46) and (47), we have uniformly in $[0, h]$

$$\sum_{|x-x_{k,n}|\leqq\delta} |\, u_{k,n}(x) \,|\, l_{k,n}^4(x) < \sum_{|x-x_{k,n}|\leqq\delta} u_{k,n}(x) l_{k,n}^4(x)$$

$$+ \sum_{k=1}^{n} l_{k,n}^4(x) < 1 + o(1) + 25.$$

This completes the proof of our lemma.

The methods used earlier, based on Lemmas 9 and 10 and Theorem IV, enable us to state the following

THEOREM V. *Let $f(x)$ be subject to the conditions of Theorem* IV. *For Laguerre abscissas, with $-\frac{3}{4} \leqq \alpha < -\frac{1}{4}$, we have uniformly in* $[0, A]$, $\lim_{n\to\infty} F_n[f] = f(x)$.

We remark, further, that for $\alpha = -\frac{1}{4}$ and $x = 0$ we can show that there exists a function $f(x)$ satisfying the conditions of Theorem IV, such that $F_n(0)$ does not converge to $f(0)$ as $n \to \infty$. The demonstration, based on the estimates for $\sum$ in Lemma 10, follows the method used with regard to Theorem II and is therefore omitted.

6. **Hermite abscissas.** As previously, we readily establish for $\omega(x) = H_n(x)$, the orthonormal Hermite polynomial, that $u_{k,n}(x) \geqq 0$ for $|x - x_{k,n}| \leqq \delta$ $(k = 1, 2, \cdots, n)$ and $|x| \leqq A$ (δ sufficiently small, and A an arbitrarily preassigned positive constant), provided n is large enough.

LEMMA 11. *For Hermite abscissas, we have uniformly for $|x| \leqq A$, as $n \to \infty$,*

$$\sum_{|x-x_{k,n}|>\delta} |u_{k,n}(x)|\, l_{k,n}^4(x) = O(n^{-\frac{1}{2}}), \qquad \sum_{|x-x_{k,n}|\leqq\delta} |u_{k,n}(x)|\, l_{k,n}^4(x) = 1 + o(1)$$

$$\left(\delta \leqq \frac{1}{8A+8}\right).$$

Here, we use ([8], p. 116; [6], pp. 210, 212; [5], p. 134),

$$(48)\quad \begin{cases} |x_{k,n}| \leqq (2n+1)^{\frac{1}{2}}, \qquad |\omega(x)| < \tau n^{-\frac{1}{4}} & \text{uniformly in } [-A, A], \\ \lambda_{k,n}\ (\equiv H_{k,n}) = \dfrac{2}{[\omega'(x_{k,n})]^2} < \tau n^{-\frac{1}{2}} & (k = 1, 2, \cdots, n). \end{cases}$$

The proof is similar to that of Lemma 8 and is therefore omitted. Proceeding precisely as in Theorem IV, we now establish

THEOREM VI. *Let $f(x)$ be continuous for all finite x and let $f(x) = O(|x|^m)$, $x \to \pm\infty$ (m arbitrarily large, fixed positive integer). Then, for the Hermite abscissas, $F_n[f] \to f(x)$, as $n \to \infty$, uniformly over any preassigned finite interval* $[-A, A]$.

REFERENCES

1. L. FEJÉR, *Über Interpolation*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen (1916), pp. 66–91.
2. L. FEJÉR, *On the characterization of some remarkable systems of points of interpolation by means of conjugate points*, American Mathematical Monthly, vol. 41(1934), pp. 1–14.
3. N. KRYLOFF AND E. STAYERMANN, *Sur quelques formules d'interpolation convergentes pour toute fonction continue*, Bulletin de la Classe des Sciences Physiques et Mathématiques de l'Académie des Sciences de l'Ukraine, vol. 1(1922), pp. 13–16.

4. E. Laguerre, *Sur les équations algébriques dont le premier membre satisfait à une équation différentielle linéaire du second ordre*, Comptes Rendus, vol. 90(1880), pp. 809–812; *Oeuvres*, vol. 1, pp. 126–132.
5. J. A. Shohat, *On interpolation*, Annals of Mathematics, (2), vol. 34(1933), pp. 130–146.
6. J. A. Shohat, *Application of orthogonal Tchebycheff polynomials to Lagrangean interpolation and to the general theory of polynomials*, Annali di Mathematica, (IV), vol. 18(1939), pp. 201–238.
7. G. Szegö, *Über gewisse Interpolationspolynome, die zu den Jacobischen und Laguerreschen Abszissen gehören*, Mathematische Zeitschrift, vol. 35(1932), pp. 579–602.
8. G. Szegö, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, vol. 23, New York, 1939.

University of Pennsylvania.

YOUNG'S SEMI-NORMAL REPRESENTATION OF THE SYMMETRIC GROUP

By R. M. Thrall

Introduction. The main purpose of this note is to give a new (shorter and more elementary) derivation of A. Young's semi-normal representation of the symmetric group. As a starting point we take the discussions by H. Weyl ([6], Chap. IV, §2, 3) and D. E. Littlewood ([4], Chap. V, especially §4).

Denote the partition $m = \lambda_1 + \cdots + \lambda_\kappa$, $\lambda_1 \geqq \cdots \geqq \lambda_\kappa > 0$ by (λ). We represent (λ) geometrically by an array of squares; λ_1 in the 1st row, $\cdots$, λ_κ in the κ-th row; the j-th squares of the rows making a column. The m squares or *fields* of the array are labelled by the numbers from 1 to m in such a way that the labels in every row increase from left to right and in every column increase from top to bottom. The array thus labelled is called a *regular Young diagram* belonging to the partition (λ).

Associated with each partition (λ) of m there is an irreducible matrix representation of the symmetric group, $\mathfrak{S}_m$, of degree m. The degree[1] $g(\lambda)$ of this representation is equal to the number of regular Young diagrams belonging to (λ). Let the label of the field in the α-th row and β-th column of a regular diagram, T, be denoted by $a(\alpha, \beta)$. If T and T' both belong to (λ) we say that

(1) T *precedes* T' if each of the fields labelled $m, m - 1, \cdots, m - r + 1$ lies in the same row in both diagrams, but the field $m - r$ lies in a lower row in T than in T'.

We enumerate the regular diagrams belonging to (λ) according to this ordering. Now number the partitions (λ) of m according to their dictionary order[2] and denote by $T(ij)$ the j-th regular Young diagram belonging to the i-th partition of m.

Corresponding to each diagram $T(ij)$ we shall define a primitive idempotent $e(ij)$ in the group $\mathfrak{K}$-ring, $\mathfrak{R}_m$, of $\mathfrak{S}_m$. [$\mathfrak{K}$ is here the field of complex numbers.]

Let $\epsilon(i) = \sum e(ij)$, summed for j from 1 to $g(\lambda^i)$. Then the two sided ideal $\epsilon(i)\mathfrak{R}_m$ of $\mathfrak{R}_m$ is a total matrix algebra $\mathfrak{A}_i = \mathfrak{A}(\lambda^i)$, of degree $g(\lambda^i)$, homomorphic with $\mathfrak{R}_m$ under the mapping $x \to x(i) = \epsilon(i)x$; and $\mathfrak{R}_m$ is the direct sum of the simple algebras $\mathfrak{A}_i$.

The next step is the choice of elements $e(ijk)$, j, $k = 1, \cdots, g(\lambda^i)$, which constitute an ordinary matrix basis ([1], p. 7) for $\mathfrak{A}_i$. In the terminology of representation theory the element x of $\mathfrak{R}_m$ is ordered to the matrix $B_i(x) =$

Received April 21, 1941.

[1] [4], Th. I, p. 68, Th. IV, p. 75; [6], Th. 7.7B, p. 213.

[2] That is, (λ) has a smaller number than (λ') if the first non-vanishing difference $\lambda_1 - \lambda_1'$, $\lambda_2 - \lambda_2'$, ... is positive.

$\| b(ijk)(x) \|$ (row index j, column index k) where the coefficients are defined by $x(i) = \sum b(ijk)(x)e(ijk)$.

With $e(ijj) = e(ij)$ and suitable choice of the non-diagonal basis elements $e(ijk)$ the following theorem holds ([7 VI], Th. IV, V, pp. 217–218). Details of this "suitable" basis (called the *orthogonal semi-normal basis*) are given below.

THEOREM I. *Let t_r be the transposition of two consecutive letters $r - 1$ and r. Then the orthogonal semi-normal matrix $B_i(t_r)$ has zero everywhere except as follows:*
(i) *If $r - 1$ and r lie in the same row of $T(ij)$, then $b(ijj)(t_r) = 1$.*
(ii) *If $r - 1$ and r lie in the same column of $T(ij)$, then $b(ijj)(t_r) = -1$.*
(iii) *If $T(ij)$ has $r - 1$ in the α-th row and β-th column, r in the γ-th row and δ-th column where $\alpha \neq \gamma$, $\beta \neq \delta$ and $T(ik)$ differs from $T(ij)$ only by exchange of $r - 1$ and r then*

$$\begin{Vmatrix} b(ijj)(t_r) & b(ijk)(t_r) \\ b(ikj)(t_r) & b(ikk)(t_r) \end{Vmatrix} = \begin{Vmatrix} -\rho & (1 - \rho^2)^{\frac{1}{2}} \\ (1 - \rho^2)^{\frac{1}{2}} & \rho \end{Vmatrix},$$

where $1/\rho = \gamma - \alpha + \beta - \delta$.

Since the transpositions t_r generate $\mathfrak{S}_m$, Theorem I provides a ready means for determining the matrices of the irreducible representations of the symmetric group.

The proof of Theorem I given below is simpler than Young's original proof. This is due primarily to formula (2) which gives a definition of the semi-normal idempotents directly in terms of the regular diagrams. Young gives no formula at all for these idempotents.

Another advantage of the present procedure is that using the properties of the semi-normal idempotents one can also obtain specific (unitary) forms for the irreducible representations of the alternation group $\mathfrak{A}_m$. This is done in §6 below.

1. **Preliminaries.** Suppose that objects are placed in the fields of a regular diagram T belonging to (λ). We shall understand by the application of the permutation

$$s: 1 \to 1', 2 \to 2', \cdots$$

to T that the object in field 1 is moved to field $1'$, the object in field 2 to field $2'$, $\cdots$. The permutation s followed by the permutation

$$t: 1' \to 1'', 2' \to 2'', \cdots$$

shall be denoted by

$$ts: 1 \to 1'', 2 \to 2'', \cdots.$$

By the *positive symmetric group* on m symbols we mean the sum of all the $m!$ permutations in $\mathfrak{S}_m$, and by the *negative symmetric group* we mean the sum of all even permutations in $\mathfrak{S}_m$ minus the sum of all the odd permutations in $\mathfrak{S}_m$.

Denote by $\mathfrak{p}_i$ the positive symmetric group on the field labels of the i-th row

of T and let $\mathfrak{p}$ be the product of $\mathfrak{p}_1, \cdots, \mathfrak{p}_s$. Let $\mathfrak{q}_j$ denote the negative symmetric group on the field labels of the j-th column of T and let $\mathfrak{q}$ be the product of all the $\mathfrak{q}_j$. Let p stand for any permutation that appears in $\mathfrak{p}$ and let q stand for any permutation that appears in $\mathfrak{q}$. Set $\delta_s = 1$ or -1 according as s is an even or an odd permutation. Evidently $q\mathfrak{q} = \delta_q \mathfrak{q}$ and $p\mathfrak{p} = \mathfrak{p}$.

It is easy to verify that Weyl's lemma 4.2A ([6], p. 122) holds good regardless of the way the fields of the diagrams involved are labelled. We rephrase his results in the form we want them here.

Lemma I. *Let $T = T(ij)$ and $T' = T(hl)$ be two regular diagrams and let s be any permutation. If* (a) $i < h$ *then there are transpositions u (a term in $\mathfrak{p} = \mathfrak{p}(T)$) and v' (a term in $\mathfrak{q}' = \mathfrak{q}(T')$) such that $su = v's$. If* (b) $i = h$, $j = l$ *then either s is of the form qp, or there are transpositions u (in $\mathfrak{p}$) and v (in $\mathfrak{q}$) such that $su = vs$.*

[Note that although nothing is claimed here when $i = h, j \neq l$ results similar to (a) can be obtained and are used in constructing Young's "natural units."[3]]

Set $\mathfrak{c} = m!\mathfrak{q}\mathfrak{p}/g(\lambda)$; $\mathfrak{c}' = m!\mathfrak{q}'\mathfrak{p}'/g(\lambda')$.

Theorem II (a) $\mathfrak{c}x\mathfrak{c}' = 0$ *for all x in $\mathfrak{R}_m$ if $i \neq i'$.*

(b) $\mathfrak{c}s\mathfrak{c} = \delta_q\mathfrak{c}$ *if* $s^{-1} = qp$

$= 0$ *if s^{-1} is not of the form qp.*

(c) *If* $x = \sum x(s)s$ *then* $\mathfrak{c}x\mathfrak{c} = \sum \delta_q x(pq)\mathfrak{c}$, *(summed for all p in $\mathfrak{p}$ and q in $\mathfrak{q}$).*

Proof. (a) The proof for the case $i > i'$ is given by Weyl.[4] Now suppose $i < i'$ and apply Lemma Ia to s^{-1} giving $s^{-1}u = v's^{-1}$ or $s = usv'$. Now $\mathfrak{p}s\mathfrak{q}' = \mathfrak{p}(usv')\mathfrak{q}' = (\mathfrak{p}u)s(v'\mathfrak{q}') = -\mathfrak{p}s\mathfrak{q}' = 0$ and so $\mathfrak{c}s\mathfrak{c}' = 0$ for any permutation s and therefore for any x in $\mathfrak{R}_m$.

(b) If s^{-1} is not of the form qp then $s^{-1}u = vs^{-1}$ and so $\mathfrak{c}s\mathfrak{c} = 0$ as in part (a). If $s^{-1} = qp$ then $\mathfrak{p}s\mathfrak{q} = (\mathfrak{p}p^{-1})s(q^{-1}\mathfrak{q}) = \delta_q\mathfrak{p}\mathfrak{q}$ and so $\mathfrak{c}s\mathfrak{c} = \delta_q\mathfrak{c}\mathfrak{c}$. We refer to Weyl's treatment ([6], pp. 124, 125) for the proof that $\mathfrak{c}\mathfrak{c} = \mathfrak{c}$. (c) is an immediate corollary to (b).

2. The semi-normal idempotents. A characteristic property of a regular diagram, T, is that deletion of the fields labelled $m, m-1, \cdots, m-k+1$ leaves a regular diagram for $m - k$. Throughout this paper we shall denote by T^* the diagram obtained from T by dropping the m-th field and shall use a star to indicate a quantity defined by T^*. Similarly T^{**} is obtained by dropping the m-th and $(m-1)$-th fields from T and double stars shall indicate quantities defined by T^{**}.

Lemma II. *If $j \neq l$ then $T^*(ij) \neq T^*(il)$.*

Proof. If m lies in the ν-th row of $T(ij)$ and the μ-th row of $T(il)$ with $\nu \neq \mu$ then $T^*(ij)$ and $T^*(il)$ belong to different partitions of $m - 1$ and so cannot be

[3] See [7 III], [7 IV], also [4], p. 75.

[4] [6], Th. 4.3D, p. 124. Note that the "$\mathfrak{c}$" here is a numerical multiple of his.

equal. If, however, $v = \mu$ then $T^*(ij) = T^*(il)$ would force $T(ij) = T(il)$ and therefore $j = l$.

We now define for each regular diagram $T(ij)$ an idempotent in $\mathfrak{R}_m$, called the *semi-normal idempotent* belonging to $T(ij)$, by the following recursive formula:

$$e(ij) = e^*(ij)c(ij)e^*(ij) \qquad (\text{for } m > 1) \tag{2}$$

and the initial condition $e(11) = 1$ for $m = 1$ (there being just one diagram $T = T(11)$ then).

THEOREM III. *The $e(ij)$ are orthogonal idempotents, i.e. $e(ij)e(hl) = \delta(ih)\delta(jl)e(ij)$ (Kronecker's δ-function).*

Proof by induction. The theorem is true for $m = 1$ and we suppose it true for all integers less than m. We establish the orthogonality. First, suppose $i = h$, but $j \neq l$. Then by Lemma II $T^*(ij) \neq T^*(il)$ and so by our induction hypothesis $e^*(ij)e^*(il) = 0$. In view of (2) this implies $e(ij)e(il) = 0$. Next, suppose $i \neq h$. Then Theorem II part (a) with $x = e^*(ij)e^*(jl)$ gives $e(ij)e(hl) = 0$.

For the proof of idempotence we recapitulate in simpler notation. T is a regular diagram with m fields. Deleting the m-th field gives a diagram T^* and deleting the last two fields from T gives a diagram T^{**}. $e = e^*ce^*$. Write e^* in the form $e^* = \sum e^*(s)s$ (summation over all s in $\mathfrak{S}_m$). Since e^* belongs to the subalgebra $\mathfrak{R}_{m-1}$ of $\mathfrak{R}_m$ we have $e^*(s) = 0$ unless s belongs to $\mathfrak{S}_{m-1}$, i.e. omits "m." By our induction hypothesis e^* is idempotent, hence: $e^* = (e^*)^3 = (e^{**}c^*e^{**})e^*(e^{**}c^*e^{**}) = e^{**}(c^*e^*c^*)e^{**} = \mu e^*$ where $\mu = \sum \delta_{q^*}e^*(p^*q^*) = 1$ (see Theorem IIc). Again, by Theorem IIc, $ee = e^*(ce^*c)e^* = \lambda e$ where $\lambda = \sum \delta_q e^*(pq)$. Now, unless both p and q omit "m," the product pq cannot lie in $\mathfrak{S}_{m-1}$. Hence, the only terms $e^*(pq)$ in the sum for λ which contribute anything are those for which p is a p^* and q is a q^*. But since T^* is a part of T every p^* is a p and every q^* is a q, and so $\lambda = \mu = 1$, which completes the proof of Theorem III.

COROLLARY I. *If $i \neq h$, then $e(ij)xe(hl) = 0$ for any x in $\mathfrak{R}_m$.*

This follows from Theorem II.

COROLLARY II. *The $e(ij)$ are primitive idempotents whose sum is* 1.

Proof. Suppose $\mathfrak{A}$ is any semi-simple $\mathfrak{K}$-algebra whose unity element, 1, is known to be the sum of h primitive orthogonal idempotents, and suppose that $e_1, \cdots, e_h$ are any mutually orthogonal idempotents in $\mathfrak{A}$. Then $1 = e_1 + \cdots + e_h$ and the e_v are all primitive. (See [1], p. 39.)

It is well known[5] that $\mathfrak{R}_m$ is the direct sum of simple matrix algebras $\mathfrak{A}_i$, one for each partition (λ^i) of m. The degree of $\mathfrak{A}_i$ is $g(\lambda^i)$, the number of regular diagrams belonging to the partition (λ^i). Hence, 1 is the sum of $g = g(\lambda^1) + g(\lambda^2) + \cdots$ primitive idempotents. Theorem III provides us with g idempotents $e(ij)$ and so the corollary follows from the above remark.

[5] [4], Chap. IV, especially Th. I, p. 68.

Let T^* be a regular diagram with $m - 1$ fields and denote by $T_1, \cdots, T_k$ the regular diagrams that can be obtained from T^* by adding one field, and suppose the corresponding semi-normal idempotents are $e^*, e_1, \cdots, e_k$. Then

COROLLARY III. $e^* = e_1 + \cdots + e_k$.

Proof. $e^* = e^*1 = e^* \sum e(ij)$. If $e(ij)$ is one of $e_1, \cdots, e_k$ then $e^* = e^*(ij)$ and $e^*e(ij) = e(ij)$. If $e(ij)$ is not one of $e_1, \cdots, e_k$ then $e^*e(ij) = e^*e^*(ij)e(ij) = 0$.

THEOREM IV. *Let* $\epsilon(i) = \sum_j e(ij)$. *Then* $\mathfrak{A}_i = \epsilon(i)\mathfrak{R}_m$ *is a simple matrix algebra of degree* $g(\lambda^i)$ *and unity element* $\epsilon(i)$; $\mathfrak{R}_m$ *is the direct sum of the* $\mathfrak{A}_i$; *and the* $\epsilon(i)$ *form a basis for the center of* $\mathfrak{R}_m$.

Proof. Let x be any element of $\mathfrak{R}_m$ and set $x(i) = \epsilon(i)x$. Then $x = 1x = x(1) + x(2) + \cdots$. If $i \neq h$ then $x(i)\epsilon(h) = \epsilon(h)x(i) = 0$ by Corollary I. Further, $x(h) = x(h)1 = \epsilon(h)x \sum_i \epsilon(i) = \epsilon(h)x\epsilon(h) = x(h)\epsilon(h) = 1x(h) = \epsilon(h)x(h)$. Therefore $x\epsilon(h) = \epsilon(h)x$, i.e. the $\epsilon(i)$ lie in the center of $\mathfrak{R}_m$. Hence, $\mathfrak{A}_i = \epsilon(i)\mathfrak{R}_m$ is a two sided ideal of $\mathfrak{R}_m$ and $\mathfrak{R}_m$ is the direct sum of the $\mathfrak{A}_i$. Since $\epsilon(i)$ is the sum of $g(\lambda^i)$ primitive idempotents the order, $n(i)$, of $\mathfrak{A}_i$ is at most $g(\lambda^i)^2$. But since the order of $\mathfrak{R}_m$ is $m! = g(\lambda^1)^2 + g(\lambda^2)^2 + \cdots = n(1) + n(2) + \cdots$ we must have $n(i) = g(\lambda^i)^2$ and so $\mathfrak{A}_i$ is a simple matrix algebra of degree $g(\lambda^i)$.

3. Special representations of the symmetric group of degree three.

The proof of Theorem I is effected by considerations which require the properties of certain special representations of the group $\mathfrak{S}$ generated by two transpositions σ and τ with a letter in common. We shall consider representations $s \to H(s)$ of $\mathfrak{S}$ by p-rowed matrices. In order that two matrices $H(\sigma)$ and $H(\tau)$ should generate a group homomorphic with $\mathfrak{S}$ it is necessary and sufficient that $H(\sigma)^2 = H(\tau)^2 = H(\sigma\tau)^3 = E_p$. This last condition is sometimes easier to apply in the equivalent form $H(\sigma\tau)^2 = H(\tau\sigma)$.

For $p = 1$ there are just two representations:

$$\mathfrak{H}_1: \; s \to \| 1 \| \qquad \text{and} \qquad \mathfrak{H}_2: \; s \to \| \delta_s \|.$$

Unless it is explicitly stated otherwise we shall use η, ζ (or η_i, ζ_i) to mean numbers such that

$$(3) \qquad 0 < \eta \leqq \tfrac{1}{2}, \qquad 0 < \zeta, \qquad \eta^2 + \zeta^2 = 1.$$

For $p = 2$ we consider the two cases

$$\mathfrak{H}_3: \; \sigma \to \begin{Vmatrix} -\eta & \zeta \\ \zeta & \eta \end{Vmatrix}, \qquad \tau \to \begin{Vmatrix} \xi_{11} & 0 \\ 0 & \xi_{22} \end{Vmatrix} \qquad \text{and}$$

$$\mathfrak{H}_4: \; \sigma \to \begin{Vmatrix} 1 & 0 \\ 0 & -1 \end{Vmatrix}, \qquad \tau \to \begin{Vmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{Vmatrix}.$$

For $\mathfrak{H}_3$ an easy computation shows that the only possibility is

$$\eta = \tfrac{1}{2}, \qquad \xi_{11} = -\xi_{22} = 1. \tag{4}$$

For $\mathfrak{H}_4$ there are two solutions:

$$-\xi_{11} = \xi_{22} = \tfrac{1}{2}; \qquad \xi_{12}\xi_{21} = \tfrac{3}{4} \quad \text{and} \tag{5}$$

$$\xi_{12} = \xi_{21} = 0; \qquad \xi_{11} = -\xi_{22} = 1. \tag{6}$$

For $p = 3$ consider

$$\mathfrak{H}_5: \quad \sigma \rightarrow \left\Vert \begin{matrix} \theta & 0 & 0 \\ 0 & -\eta & \zeta \\ 0 & \zeta & \eta \end{matrix} \right\Vert, \quad \text{where } \theta = \pm 1; \qquad \tau \rightarrow \left\Vert \begin{matrix} \xi_{11} & \xi_{12} & 0 \\ \xi_{21} & \xi_{22} & 0 \\ 0 & 0 & \xi_{33} \end{matrix} \right\Vert.$$

Then $(\sigma\tau)^2 = (\tau\sigma)$ gives

$$\left\Vert \begin{matrix} \xi_{11}^2 - \eta\theta\xi_{12}\xi_{21} & \xi_{11}\xi_{12} - \eta\theta\xi_{12}\xi_{22} & \zeta\theta\xi_{12}\xi_{33} \\ \xi_{21}(-\eta\theta\xi_{11} + \eta^2\xi_{22} + \zeta^2\xi_{33}) & -\eta\theta\xi_{12}\xi_{21} + \eta^2\xi_{22} + (1 - \eta^2)\xi_{22}\xi_{33} & \eta\zeta\xi_{33}(-\xi_{22} + \xi_{33}) \\ \zeta\xi_{21}(\theta\xi_{11} - \eta\xi_{22} + \eta\xi_{33}) & \zeta(\theta\xi_{12}\xi_{21} - \eta\xi_{22}^2 + \eta\xi_{22}\xi_{33}) & (1 - \eta^2)\xi_{22}\xi_{33} + \eta^2\xi_{33}^2 \end{matrix} \right\Vert$$

$$= \left\Vert \begin{matrix} \theta\xi_{11} & -\eta\xi_{12} & \zeta\xi_{12} \\ \theta\xi_{21} & -\eta\xi_{22} & \zeta\xi_{22} \\ 0 & \zeta\xi_{33} & \eta\xi_{33} \end{matrix} \right\Vert.$$

Let (α, β) denote the equation obtained by equating elements in the α-th row and β-th column of the matrices representing $(\sigma\tau)^2$ and $\tau\sigma$. First suppose $\xi = \xi_{12}\xi_{21} \neq 0$. From (1, 3) we get $\xi_{33} = \theta$ and then from (3, 1), (2, 3), (1, 1) $\theta\xi_{11} - \eta\xi_{22} = -\eta\theta$, $\eta\theta(-\xi_{22} + \theta) = \xi_{22}$, $\xi_{11}^2 - \eta\theta\xi_{12}\xi_{21} = \theta\xi_{11}$ we get

$$-\xi_{11} = \xi_{22} = \eta/(1 + \eta\theta); \qquad \xi_{12}\xi_{21} = 1 - \xi_{22}^2; \qquad \xi_{33} = \theta. \tag{7}$$

If $\xi = 0$ (1, 1) gives $\xi_{11} = \theta$. $\tau^2 = 1$ requires $\xi_{22}^2 = \xi_{33}^2 = 1$ which with (2, 3) leads to

$$\eta = \tfrac{1}{2}; \qquad \xi_{11} = \theta; \qquad \xi_{22} = -\xi_{33} = 1; \qquad \xi_{12} = \xi_{21} = 0. \tag{8}$$

Finally we consider $p = 6$ with

$$\mathfrak{H}_6: \quad \sigma \rightarrow \left\Vert \begin{matrix} -\eta_1 & \zeta_1 & & & & \\ \zeta_1 & \eta_1 & & & & \\ & & -\eta_2 & \zeta_2 & & \\ & & \zeta_2 & \eta_2 & & \\ & & & & -\eta_3 & \zeta_3 \\ & & & & \zeta_3 & \eta_3 \end{matrix} \right\Vert, \qquad \tau \rightarrow \left\Vert \begin{matrix} \xi_{11} & & \xi_{13} & & & \\ & \xi_{22} & & & \xi_{25} & \\ \xi_{31} & & \xi_{33} & & & \\ & & & \xi_{44} & & \xi_{46} \\ & \xi_{52} & & & \xi_{55} & \\ & & & \xi_{64} & & \xi_{66} \end{matrix} \right\Vert$$

(blank spaces being filled with zeros) with the additional relation $1/\eta_2 = 1/\eta_1 + 1/\eta_3$. We now consider the relations obtained by equating the matrices representing $\tau\sigma$ and $(\sigma\tau)^2$.

First case, $\xi = \xi_{13}\xi_{31}\xi_{25}\xi_{52}\xi_{46}\xi_{64} \neq 0$. Then $\tau^2 = 1$ requires $-\xi_{11} = \xi_{33}$, $-\xi_{22} = \xi_{55}$, $-\xi_{44} = \xi_{66}$, $\xi_{13}\xi_{31} = 1 - \xi_{33}^2$, $\xi_{25}\xi_{52} = 1 - \xi_{55}^2$, $\xi_{46}\xi_{64} = 1 - \xi_{66}^2$. Now the equations (3, 1), (4, 1), (6, 2) together with $\xi \neq 0$ and the relation connecting the η_i give the single solution

$$\xi_{33} = \eta_3, \qquad \xi_{66} = \eta_1, \qquad \xi_{55} = \eta_2. \tag{9}$$

Second case, $\xi = 0$. From $\tau^2 = 1$ we get six implications of the form $\xi_{ij} = 0$ implies $\xi_{ii}\xi_{jj} \neq 0$ with $i, j = 1, 3; 3, 1; 2, 5; 5, 2; 4, 6; 6, 4$. These inequalities together with equations (6, 1), (4, 2), (5, 3), (2, 4), (3, 5), (1, 6) show that for ξ to vanish each of its six factors must be zero. But then by (4) we should have to have each $\eta_i = \frac{1}{2}$ whereas by (3) and the relation between the η_i we have $\eta_2 \leqq \frac{1}{4}$, thus ruling out this case.

For later use we note also that in case $\xi_{13} = \xi_{31} = \zeta_3$, $\xi_{25} = \xi_{52} = \zeta_2$ then (6, 1): $-\zeta_1\zeta_3\eta_3\xi_{52} + \eta_3\zeta_2\xi_{31}\xi_{64} = 0$ requires

$$\xi_{46} = \xi_{64} = \zeta_1, \tag{10}$$

or in other words requires that $\mathfrak{H}_6$ be orthogonal.

4. **Proof of Theorem I: first part.** We shall suppose that Theorem I has been verified by direct computation for the symmetric groups of degree 4 or less. Proceeding by induction we suppose that it is already established for the symmetric groups of degree less than m and proceed to the proof for degree m. The proof is divided into two parts, a "qualitative" part and a "quantitative" part. In this section we show (always based upon our induction hypothesis) that the theorem is true for transpositions t_r, for $r < m$, and that the matrix representing $\tau = t_m$ has zeros in the places required. Then in the next section we complete the proof by showing that the non-diagonal matrix units can be chosen so that the matrix for τ takes the form specified by Theorem I. We mention here that by a slightly different choice of non-diagonal matrix units a rational representation can be obtained ([7 VI], p. 217).

Let (λ) be the partition of m defined by $m = n_1\lambda_1 + \cdots + n_s\lambda_s$, $\lambda_1 > \lambda_2 > \cdots > \lambda_s > 0$, and denote by $D_i(\lambda)$ the partition of $m - 1$ defined by $m - 1 = n_1\lambda_1 + \cdots + (n_i - 1)\lambda_i + (\lambda_i - 1) + \cdots + n_s\lambda_s$. Let $T_1, T_2, \cdots$ be the regular diagrams belonging to (λ) ordered according to (1); let $e_1, e_2, \cdots$ be the corresponding semi-normal idempotents. Denote by $\mathfrak{A}$ the simple matrix algebra $\epsilon\mathfrak{R}_m$ where $\epsilon = e_1 + e_2 + \cdots$. Let $\mathfrak{T}(i)$ be the set of diagrams T_j which have m in the last row of length λ_i. We observe that omission of the m-th field from each diagram in $\mathfrak{T}(i)$ gives the complete set of regular diagrams belonging to the partition $D_i(\lambda)$ of $m - 1$.

LEMMA III. *If T_j and T_k belong to different sets $\mathfrak{T}(i)$ and x lies in $\mathfrak{R}_{m-1}$ then $e_jxe_k = 0$.*

Proof. We have $e_v = e_v^*c_ve_v^*$. If T_j and T_k lie in different sets $\mathfrak{T}(i)$ then e_j^* and e_k^* arise from different partitions $D_i(\lambda)$ of $m - 1$ and so $e_j^*xe_k^* = 0$ by Corollary I (for $m - 1$).

Let ϵ_i^* denote the sum of the e_j^* for which T_j lies in $\mathfrak{T}(i)$. If T_j is in $\mathfrak{T}(i)$, $\epsilon e_{jj}^* = e_j$ by Corollary III and Theorem III. Thus $\epsilon\epsilon_i^*$ is the sum of all the e_j for which T_j is in $\mathfrak{T}(i)$. Adding we obtain $\epsilon\epsilon_\kappa^* + \cdots + \epsilon\epsilon_1^* = \epsilon$. Let $\mathfrak{A}_i^* = \epsilon\epsilon_i^*\mathfrak{R}_{m-1}$.

Lemma IV. *$\epsilon\mathfrak{R}_{m-1}$ is the direct sum of the algebras $\mathfrak{A}_i^*$, and $\mathfrak{A}_i^*$ is a simple matrix algebra which is the image of $\mathfrak{R}_{m-1}$ under the homomorphic mapping $x \rightarrow \epsilon\epsilon_i^* x$.*

Proof. The first statement is a corollary to Lemma III. By Theorem IV (for $m - 1$) $\epsilon_i^*\mathfrak{R}_{m-1}$ is a simple matrix algebra homomorphic to $\mathfrak{R}_{m-1}$ under the mapping $x \rightarrow \epsilon_i^* x$. Now since ϵ lies in the center of $\mathfrak{R}_m$ the mapping $\epsilon_i^* x \rightarrow \epsilon\epsilon_i^* x$ is a homomorphism. Further, $\epsilon\epsilon_i^* \neq 0$ implies $\mathfrak{A}_i^* \neq 0$. But the only non-zero homomorphic image of a simple matrix algebra is a simple matrix algebra isomorphic to it.

Let T_j belong to $\mathfrak{T}(i)$. Then $e_j = \epsilon\epsilon_i^* e_j^* = \epsilon e_j^*$ and so in the above indicated isomorphism between $\mathfrak{A}_i^*$ and $\epsilon_i^*\mathfrak{R}_{m-1}$ the idempotent e_j maps into a diagonal element of $\epsilon_i^*\mathfrak{R}_{m-1}$. If T_k also belongs to $\mathfrak{T}(i)$ we define e_{jk} as the image, ϵe_{jk}^*, of the corresponding matrix basis element e_{jk}^* of $\epsilon_i^*\mathfrak{R}_{m-1}$. By our induction hypothesis Theorem I is true for $m - 1$ and therefore applies to each $\epsilon_i^*\mathfrak{R}_{m-1}$, $i = \kappa, \cdots, 1$. Thus with the above choice of matrix basis elements e_{jk} for T_j and T_k in the same set $\mathfrak{T}(i)$ and any choice[6] of the remaining matrix basis elements we have[7]

Lemma V. *Let x lie in $\mathfrak{R}_{m-1}$. Then the matrix $B(x) = \| b(jk\wr x) \|$ defined by $\epsilon x = \sum b(jk\wr x)e_{jk}$ has zeros outside blocks, $B_i(x)$, $i = \kappa, \cdots, 1$, along the main diagonal. The block $B_i(x)$ is made up of those elements $b(jk\wr x)$ for which T_j and T_k lie in $\mathfrak{T}(i)$. The block $B_i(x)$ is a matrix identical with the semi-normal matrix for x in the irreducible representation of $\mathfrak{R}_{m-1}$ defined by the partition $D_i(\lambda)$ of $m - 1$.*

Observe that, in particular, Lemma V for $x = t_r$ implies Theorem I for $r < m$.

Denote by $D_{ih}(\lambda)$, $i \neq h$, the partition of $m - 2$ defined by

$$m - 2 = \cdots + (n_i - 1)\lambda_i + (\lambda_i - 1) + \cdots + (n_h - 1)\lambda_h + (\lambda_h - 1) + \cdots .$$

Observe that $D_{ih}(\lambda) = D_{hi}(\lambda)$. If $(\lambda_i - 1) > \lambda_{i-1}$ denote by $D_i\leftarrow(\lambda)$ the partition of $m - 2$ defined by

$$m - 2 = \cdots + (n_i - 1)\lambda_i + (\lambda_i - 2) + \cdots;$$

and if $n_i > 1$ denote by $D_i \uparrow (\lambda)$ the partition defined by

$$m - 2 = \cdots + (n_i - 2)\lambda_i + 2(\lambda_i - 1) + \cdots .$$

Let $\mathfrak{T}(ih)$ be the set of diagrams, T_j, in $\mathfrak{T}(i)$ with the $(m - 1)$-st field in the last row of length λ_h. If we order the diagrams in a set by (1) the j-th member

[6] That is, any choice consistent with the multiplication rules for an ordinary matrix basis. In §5 we make this choice specific.

[7] [7 VI], p. 217. The rest of the present section is along the lines of Young's argument.

of $\mathfrak{T}(ih)$ becomes the j-th member of $\mathfrak{T}(hi)$ by exchange of the fields m and $m - 1$.

Similarly let $\mathfrak{T}(i\leftarrow)[\mathfrak{T}(i\uparrow)]$ be the set of diagrams in $\mathfrak{T}(i)$ with the fields $m - 1$ and m in the same row [column].

Now by Lemma V, for $m - 1$, the matrix $B_i(x)$, for x in $\mathfrak{R}_{m-2}$, has zeros outside blocks $B_{ii}(x), \cdots, B_{ii+1}(x), B_i\leftarrow(x), B_i\uparrow(x), \cdots, B_{i1}(x)$ down the main diagonal. The block $B_{ih}(x)$ is made up of those elements $b(jk\between x)$ for which T_j and T_k lie in $\mathfrak{T}(ih)$. $x \to B_{ih}(x)$ is the semi-normal representation, $\mathfrak{B}_{ih}$ of $\mathfrak{R}_{m-2}$ defined by the partition $D_{ih}(\lambda)$. Similar statements hold for the matrices $B_i\leftarrow(x)$ and $B_i\uparrow(x)$ in case they occur. From $D_{ih}(\lambda) = D_{hi}(\lambda)$ we infer that $B_{ih}(x) = B_{hi}(x)$; and since the partitions $D_{ih}(\lambda)$, $D_j\leftarrow(\lambda)$, $D_k\uparrow(\lambda)$ are otherwise distinct no further equivalences between the representations $\mathfrak{B}_{ih}$, $\mathfrak{B}_j\leftarrow$, $\mathfrak{B}_k\uparrow$ are possible.

Denote by $C_{ih}(C_i\leftarrow, C_i\uparrow)$ the $g(\lambda)$-rowed square matrix with zeros everywhere save for 1's down the main diagonal of the block occupied by $\mathfrak{B}_{ih}(\mathfrak{B}_i\leftarrow, \mathfrak{B}_i\uparrow)$, and by C'_{ih} the matrix with zeros everywhere except for 1's down the main diagonal of the block formed by the intersection of the rows of $\mathfrak{B}_{ih}$ with the columns of $\mathfrak{B}_{hi}$. It follows from the above arguments and the theory of semi-simple matrix algebras[8] that we have

Lemma VI. *Let C be any $\mathfrak{R}$-matrix commutative with $B(x)$ for every x in $\mathfrak{R}_{m-2}$. Then there exist multipliers ρ_{ih}, ρ'_{ih}, $\rho_i\leftarrow$, $\rho_i\uparrow$ in $\mathfrak{R}$ such that*

$$C = \sum_{i\neq h} \rho_{ih} C_{ih} + \rho'_{ih} C'_{ih} + \sum_i \rho_i\leftarrow C_i\leftarrow + \rho_i\uparrow C_i\uparrow. \tag{11}$$

Now τ is commutative with any x in $\mathfrak{R}_{m-2}$ and so $B(\tau)$ is of the form (11), which puts zeros in just the places demanded by Theorem I.

For the next section we shall need to interpret the foregoing results in terms of the relations between τ and the matrix units e_{jk}. We recall that $B(\tau)$ is defined by $e_j\tau e_k = b(jk\between\tau)e_{jk}$.

Lemma VII. *$b(jk\between\tau) = 0$ unless $j = k$ or T_j and T_k differ by exchange of $m - 1$ and m. If T_j and $T_{j'}$ lie in the same set $\mathfrak{T}(ih)$, $\mathfrak{T}(i\leftarrow)$, or $\mathfrak{T}(i\uparrow)$ then $b(jj\between\tau) = b(j'j'\between\tau)$. If T_j and $T_{j'}$ lie in the same set $\mathfrak{T}(ih)$ and are replaced by T_k and $T_{k'}$ upon exchange of $m - 1$ and m then $b(jk\between\tau) = b(j'k'\between\tau)$.*

In other words we need make only one computation for each position of $m - 1$ and m and in that computation may place the first $m - 2$ fields in arbitrary positions (subject of course to regularity). We use this fact freely in the next section.

5. Proof of Theorem I: second part.

Lemma VIII. *Let T_j be a regular diagram belonging to the partition (λ) of m and with $r - 1 = a(\alpha, \beta)$, $r = a(\gamma, \delta)$. Set $(r - 1, r) = t_r$.*

(i) *If $\alpha = \gamma$ then $e_jt_re_j = e_j$.*

(ii) *If $\beta = \delta$ then $e_jt_re_j = -e_j$.*

[8] See [6], Th. 3.5B, p. 95.

(iii) *If* $\alpha \neq \gamma$, $\beta \neq \delta$ *let* T_k *be obtained from* T_j *by exchange of* $r - 1$ *and* r. *Then* $e_j t_r e_j = -\rho e_j$; $e_k t_r e_k = \rho e_k$; $e_j t_r e_k t_r e_j = (1 - \rho^2)e_j$; $e_k t_r e_j t_r e_k = (1 - \rho^2)e_k$ *where* $1/\rho = \gamma - \alpha + \beta - \delta$. We call $1/\rho$ the *axial projection* of the fields $a(\alpha, \beta)$, $a(\gamma, \delta)$.

Proof. For $r < m$ Lemma VIII is a consequence of Lemma V and our induction hypothesis. Set $t_{m-1} = \sigma$. σ and τ generate the symmetric group, $\mathfrak{S}_3'$, on $m - 2, m - 1, m$. We establish Lemma VIII for $\tau(= t_m)$ by requiring that $B(\sigma\tau)^2 = B(\tau\sigma)$.

We may simplify the computations by dividing the diagrams T_v into classes. T_j and T_k are said to belong to the same *class*, $\mathfrak{C}$, if the diagrams obtained from them by deleting the last three fields are identical. Let $T_{v(1)}, \cdots, T_{v(p)}$ be the diagrams in a class $\mathfrak{C}$, where $v(1) < \cdots < v(p)$. Suppose $m - 2 = a(\alpha_1, \beta_1)$, $m - 1 = a(\alpha_2, \beta_2)$, $m = a(\alpha_3, \beta_3)$ in $T_{v(1)}$. Then by (1)

$$(12) \qquad \alpha_1 \leqq \alpha_2 \leqq \alpha_3 .$$

The number, p, of diagrams in $\mathfrak{C}$ is 1, 2, 3 or 6 according to the distribution of equalities in (12), and between the β's.

Consider the p-rowed matrices $B_{\mathfrak{C}}(x) = \| x(jk) \|$ defined by $x(jk) = b(v(j)v(k) \mid x)$. It follows from Lemmas V and VII that the matrix $B(x)$ representing any element x in $\mathfrak{R}_3'$ (the group ring of $\mathfrak{S}_3'$) will have $b(jk \mid x) = 0$ unless T_j and T_k belong to the same class $\mathfrak{C}$. Hence by rearranging rows (and corresponding columns) we can bring $B(x)$ into a form with blocks $B_{\mathfrak{C}}(x)$ along the main diagonal and 0 elsewhere. Therefore the matrices $B_{\mathfrak{C}}(x)$ afford a representation $\mathfrak{B}_{\mathfrak{C}}$ of $\mathfrak{R}_3'$ and $B(\sigma\tau)^2 = B(\tau\sigma)$ is equivalent to $B_{\mathfrak{C}}(\sigma\tau)^2 = B_{\mathfrak{C}}(\tau\sigma)$ for each class $\mathfrak{C}$.

Consider the following operations on a representation $\mathfrak{B}_{\mathfrak{C}}$ of $\mathfrak{S}_3'$:

(13) (i) applying a permutation to the rows and columns of each $B_{\mathfrak{C}}(x)$;
(ii) changing signs in the i-th row and then in the i-th column;
(iii) replacing $B_{\mathfrak{C}}(s)$ by $\delta_s B_{\mathfrak{C}}(s)$.

We shall call these *admissible operations.*

Lemma IX. *Any representation* $\mathfrak{B}_{\mathfrak{C}}$ *of* $\mathfrak{S}_3'$ *can be sent into one of the special representations* $\mathfrak{H}_v$ *by means of admissible operations.*

The proof of Lemma IX is divided into cases according to the positions of the last three fields in the diagrams of $\mathfrak{C}$. We make use of the parts of Lemma VIII that have already been established to determine $B_{\mathfrak{C}}(\sigma)$ and the form of $B_{\mathfrak{C}}(\tau)$.

If the fields $m - 2, m - 1, m$ lie in a row (column), then $\mathfrak{B}_{\mathfrak{C}} = \mathfrak{H}_1(\mathfrak{H}_2)$.

If they lie in three rows and three columns, then $\mathfrak{B}_{\mathfrak{C}} = \mathfrak{H}_6$, if we choose $1/\eta_1 = \beta_1 - \beta_2 + \alpha_2 - \alpha_1$, $1/\eta_2 = \beta_1 - \beta_3 + \alpha_3 - \alpha_1$, $1/\eta_3 = \beta_2 - \beta_3 + \alpha_3 - \alpha_2$ (and hence $1/\eta_2 = 1/\eta_1 + 1/\eta_3$).

If they lie in two rows and two columns there are two possibilities:

$$(14) \qquad \alpha_1 + 1 = \alpha_2 = \alpha_3, \qquad \beta_1 = \beta_2 + 1 = \beta_3 \quad \text{which gives}$$

$$\mathfrak{B}_{\mathfrak{C}} = \mathfrak{H}_3 \quad \text{with} \quad \eta = \tfrac{1}{2}; \qquad \text{and}$$

(15) $\alpha_1 = \alpha_2 = \alpha_3 - 1, \quad \beta_1 - 1 = \beta_2 = \beta_3$ which gives $\mathfrak{B}_\varepsilon = \mathfrak{H}_4$.

If they lie in three rows and two columns or three columns and two rows there are four possibilities, all of which lead to $\mathfrak{H}_5$:

$$(16) \qquad \alpha_1 < \alpha_2 = \alpha_3, \qquad \beta_1 > \beta_2 + 1 = \beta_3.$$

Take $1/\eta = \beta_2 - \beta_3 + \alpha_3 - \alpha_2$ and $\theta = -1$. To obtain $\mathfrak{H}_5$ from $\mathfrak{B}_\varepsilon$ interchange the first and third rows and columns, change signs in the second row and column, and apply 13-(iii). Note that $\eta < \frac{1}{2}$ so that solution (8) is ruled out.

$$(17) \qquad \alpha_1 = \alpha_2 < \alpha_3, \qquad \beta_1 - 1 = \beta_2 > \beta_3.$$

Take $1/\eta = \beta_1 - \beta_3 + \alpha_3 - \alpha_1$, $\theta = 1$ and we have $\mathfrak{B}_\varepsilon = \mathfrak{H}_5$. Solution (8) can occur only if $\eta = \frac{1}{2}$; i.e. $\alpha_3 = \alpha_2 + 1$, $\beta_2 = \beta_3 + 1$.

We omit the details for the other two cases:

$$(18) \qquad \alpha_1 < \alpha_2 = \alpha_3 - 1, \qquad \beta_1 - 1 = \beta_2 > \beta_3 \quad \text{and}$$

$$(19) \qquad \alpha_1 + 1 = \alpha_2 < \alpha_3, \qquad \beta_1 = \beta_2 > \beta_3.$$

In (19) solution (8) cannot occur, and in (18) it can occur only if $\alpha_1 = \alpha_2 - 1$, $\beta_1 = \beta_2 + 1$.

This completes the proof of Lemma IX. Furthermore, except for (6) and (8), the solutions for $H_\nu(\tau)$ in §3, and hence for any representations obtained from them by admissible operations, satisfy Lemma VIII; hence to complete its proof we need only show that solution (6) in case (15) and solution (8) in cases (17) and 18) cannot occur.

Suppose then that $\mathfrak{C}$ is in case (15). If $\beta_3 > 1$ consider a diagram T with $m - 2$ in the position $(\alpha_3, \beta_3 - 1)$ and $m - 1$, m as in $T_{\nu(1)}$. The class $\mathfrak{C}'$ containing T is of type (16). From Lemma VII and a comparison of (6) and (7) we see that (6) is not possible here. If $\beta_3 = 1$ we have from $m > 4$ that $\alpha_1 > 1$ and so may rule out (6) by comparison with a class $\mathfrak{C}'$ containing a diagram with $m - 1$, m in the same positions as $T_{\nu(1)}$ and $m - 2$ in the place $(\alpha_1 - 1, \beta_1')$.

Next consider (17) and suppose that solution (7) actually holds, which implies among other things that $e_{\nu(1)}\tau e_{\nu(1)} = e_{\nu(1)}$. If $\beta_3 > 1$ take a diagram T with m, $m - 1$ as in $T_{\nu(1)}$ and $m - 2$ as in the position $(\alpha_3, \beta_3 - 1)$. The class $\mathfrak{C}'$ containing T has three diagrams $T'_{\nu(1)}$, $T'_{\nu(2)}[= T]$, $T'_{\nu(3)}$. $\mathfrak{C}'$ is in case (16) and so $e'_{\nu(2)}\tau e'_{\nu(2)} = -e'_{\nu(2)}/2$. But since $1 \neq -1/2$ this contradicts Lemma VII. If $\beta_3 = 1$, then from $m > 4$ we conclude $\alpha_1 > 1$ and proceed as above for case (15). The proof of case (18) is the same, upon interchange of the roles of rows and columns of the related diagrams. This completes the proof of Lemma VIII.

Choose diagrams $T_{\nu(i)}$, $i = \kappa - 1, \cdots, 1$ in $\mathfrak{T}(\kappa i)$ and suppose $T_{\mu(i)}$ obtained from $T_{\nu(i)}$ by exchange of $m - 1$ and m. Let $1/\eta_i$ be the axial projection of the fields $m - 1$ and m in $T_{\nu(i)}$. We now complete the choice of a matrix basis for the simple algebra $\mathfrak{A}$ of which $\mathfrak{B}$ is a faithful representation. The diagonal units are the semi-normal idempotents. If T_j and T_k lie in the same set $\mathfrak{T}(i)$ we define e_{jk} as in Lemma V. Set $e_{\nu(i)\mu(i)} = e_{\nu(i)}\tau e_{\mu(i)}/\zeta_i$, and define the re-

maining basis elements by the multiplication rules $e_{jk}e_{j'k'} = \delta_{kj'}e_{jk'}$. In terms of this basis, Lemmas VII and VIII give $b(jk)(\tau) = \zeta_i = b(kj)(\tau)$ if T_j lies in $\mathfrak{T}(\kappa i)$ and T_k is obtained from T_j by exchange of $m-1, m$. If finally $h < i < \kappa$ then consider a diagram T in $\mathfrak{T}(ih)$ which has $m - 2$ as the final field in its last row. The class $\mathfrak{C}$ containing T has six diagrams. Arranging these in the usual order, and making use of the symmetry properties of $B(\tau)$ already established we have $\xi_{13} = \xi_{31} = \zeta_i$, $\xi_{25} = \xi_{52} = \zeta_h$ and therefore by (10) $0 < \xi_{64} = \xi_{46}$ so that every $B_{\mathfrak{C}}(\tau)$ is orthogonal. Therefore $B(\tau)$ is orthogonal, and Theorem I is proved.

6. **The irreducible representations of the alternating group.** Let (λ) be a partition of m and $\mathfrak{B}$: $s \to B(s)$ be the corresponding semi-normal orthogonal representation of $\mathfrak{S}_m$. The matrices $B(s)$ for s in the alternating group, $\mathfrak{A}_m$, constitute a representation of $\mathfrak{A}_m$ which we shall denote by $\mathfrak{B}\dagger$. If (λ) is associated[9] with (μ) then $\mathfrak{B}\dagger$ is equivalent to the representation of $\mathfrak{A}_m$ defined by (μ). If $(\lambda) \neq (\mu)$, $\mathfrak{B}\dagger$ is an irreducible representation of $\mathfrak{A}_m$. If $(\lambda) = (\mu)$, i.e. (λ) is self associated, $\mathfrak{B}\dagger$ splits into two inequivalent (but conjugate)[10] irreducible representations of $\mathfrak{A}_m$ which we shall denote by $\mathfrak{B}\dagger_i : s \to A_i(s)$, $i = 1, 2$. With the aid of the preceding results for the symmetric group we can obtain a specific form for this splitting, indeed one such that the $\mathfrak{B}\dagger_i$ will be unitary, and will thus have determined specific unitary forms for all the irreducible representations of $\mathfrak{A}_m$.

Denote by $T_1, \cdots, T_g$ $(g = g(\lambda))$ the regular diagrams belonging to (λ) and by $T'_1, \cdots, T'_g$ the regular diagrams belonging to the associated partition (μ), both sets being ordered according to (1). Denote by $T\dagger_j$ the diagram obtained from T_j by interchanging rows and columns. $T\dagger_j$ is evidently a regular diagram belonging to (μ). It follows immediately from (1) that if $j < h$ then $T\dagger_h$ precedes $T\dagger_j$, thus giving the following lemma:

LEMMA X. $T\dagger_j = T'_{g-j+1}$.

Let $\mathfrak{A}$, $\mathfrak{A}'$ be the simple invariant subalgebras of $\mathfrak{R}_m$ defined by (λ) and (μ) respectively. If $x = \sum x(s)s$ is in $\mathfrak{A}$ then $x\dagger = \sum \delta_s x(s)s$ is in $\mathfrak{A}'$ and the mapping $x \to x\dagger$ is[11] an isomorphism between $\mathfrak{A}$ and $\mathfrak{A}'$. Hence if $\epsilon = \sum \epsilon(s)s$ is the unity element of $\mathfrak{A}$, then $\epsilon\dagger$ is the unity element ϵ' of $\mathfrak{A}'$.

LEMMA XI. *Let* $e_j = \sum e_j(s)s$ *be the semi-normal idempotent defined by* T_j. *Then* $e\dagger_j = \sum \delta_s e_j(s)s$ *is the semi-normal idempotent* e'_{g-j+1} *defined by* $T\dagger_j$.

Proof by induction on m. As we saw in the proof of Lemma IV, $e_j = \epsilon e^*_j$ and $e'_{g-j+1} = \epsilon' e^{*\prime}_{g-j+1}$. Apply our induction hypothesis $e^{*\prime}_{g-j+1} = \sum \delta_s e^*_j(s)s$, and $e\dagger_j = e'_{g-j+1}$ follows by a direct computation.

COROLLARY IV. *If* (λ) *is self associated* $e\dagger_j = e_{g-j+1}$.

[9] [5], p. 120. See Chap. VI for other results quoted in this paragraph.
[10] See [2], Th. I, p. 534.
[11] See [4], Chap. V, §4.

LEMMA XII. *If* (λ) *is self associated, then*
(i) $b(jj\between s) = \delta_s b(g-j+1, g-j+1\between s)$, *and*
(ii) $b(jk\between t_r) = b(g-j+1, g-k+1\between t_r)$, $j \neq k$.

Proof. (i) $b(g-j+1, g-j+1\between s)e_{g-j+1} = e\dagger_j s e\dagger_j = \sum_{u,v}\delta_{uv}e_j(u)e_j(v)usv = \sum_{u,t}\delta_s\delta_t e_j(u)e_j(s^{-1}u^{-1}t)t = \delta_s b(jj\between s)e_{g-j+1}$. (ii) follows from Theorem I and Corollary IV.

Let $\theta_v = 1$ or -1 according as the permutation which sends T_1 into T_v is even or odd, and set

$$C = \left\| \begin{matrix} 0 & & & & \theta_1 \\ & & & \theta_2 & \\ & & \cdot & & \\ & \cdot & & & \\ \theta_g & & & & 0 \end{matrix} \right\|.$$

Note that $C^2 = \theta_g E_g$.

THEOREM V. $C^{-1}B(s)C = \delta_s B(s)$ *for every* s *in* $\mathfrak{S}_m$.

Proof. It is sufficient to establish the theorem for the transpositions t_r. Set $A = C^{-1}B(t_r)C$. Then $a_{jk} = \theta_j\theta_k b(g-j+1, g-k+1\between t_r)$. If $b(jk\between t_r) \neq 0$ for $j \neq k$, then T_j is sent into T_k by the transposition t_r so that $\theta_j\theta_k = -1$. Now the theorem follows from Lemma XII.

COROLLARY V. *For* s *in* $\mathfrak{A}_m$, $B(s)C = CB(s)$.

Since $\mathfrak{A}_m$ is of index 2 in $\mathfrak{S}_m$ the commutator algebra of the representation $\mathfrak{B}\dagger$ must be generated by E_g and C.[12]

Since (λ) is self associated g is even, say $g = 2f$. Observe that if (i) $\theta_g = 1$, then C has f eigenvalues 1 and f eigenvalues -1 and if (ii) $\theta_g = -1$, then C has f eigenvalues i and f eigenvalues $-i$. The sign of θ_g can be determined in the following way. Set $p_j = 2(\lambda_j - j) - 1$, $j = 1, 2, \cdots, r$; r being uniquely defined by $m = p_1 + \cdots + p_r$ (see [4], p. 272). T_1 and T_g have r fields with the same label and hence from $T_g = T\dagger_1$ we conclude[13] that $\theta_g = i^{m-r}$.

LEMMA XIII. *Let*

$$Q = \left\| \begin{matrix} a_1 & & & & & b_1 \\ & \ddots & & & \cdot^{\cdot^{\cdot}} & \\ & & a_f & b_f & & \\ & & \bar{b}_f & -a_f & & \\ & \cdot^{\cdot^{\cdot}} & & & \ddots & \\ \bar{b}_1 & & & & & -a_1 \end{matrix} \right\|,$$

where the dash indicates conjugate imaginary and where if (i) $\theta_g = 1$; $\theta_j a_j = b_j = \sqrt{2}/2$; (ii) $\theta_g = -1$; $a_j = -i\theta_j b_j = \sqrt{2}/2$. *Then if* (i) $QCQ = \left\| \begin{matrix} E_f & 0 \\ 0 & -E_f \end{matrix} \right\|$; *and if* (ii) $QCQ = \left\| \begin{matrix} iE_f & 0 \\ 0 & -iE_f \end{matrix} \right\|$.

[12] See [2], Th. III, p. 538.
[13] Compare with [3], p. 310.

Proof by direct computation. Observe that Q is unitary and of order two.

The decomposition of a semi-simple matric algebra carries with it the decomposition of its commutator algebra. Hence

THEOREM VI. *Let s belong to $\mathfrak{A}_m$. Then*

$$QB(s)Q = \left\| \begin{matrix} A_1(s) & 0 \\ 0 & A_2(s) \end{matrix} \right\|,$$

and $s \to A_i(s)$ $i = 1, 2$ is an irreducible unitary representation of $\mathfrak{A}_m$. If $m - r$ is divisible by 4, $A_i(s)$ is real; if $m - r$ is not divisible by 4, some $A_i(s)$ have complex elements. If s is an odd permutation of $\mathfrak{S}_m$, then $QB(s)Q$ is of the form $\left\| \begin{matrix} 0 & * \\ * & 0 \end{matrix} \right\|$.

Only the last statement requires proof and it will be established if shown for a single odd permutation. It is readily verified for the transposition t_2 (or indeed for any t_v).

Several other properties of the representations $\mathfrak{B}\dagger_i$ follow upon closer study of the form of the matrices $QB(t_v)Q$. Of these we state, without proof, the following: suppose $m - r$ not divisible by 4. Then if (i) $p_r > 1$ and s is an even permutation which omits m, then $A_i(s)$ is real; and if (ii) $p_r = 1$, $A_i(s)$ is real if s omits m and $m - 1$.

BIBLIOGRAPHY

1. ALBERT, A. A., *Structure of Algebras*, American Mathematical Society, New York, 1939.
2. CLIFFORD, A. H., *Representations induced in an invariant subgroup*, Annals of Mathematics, vol. 38(1937), pp. 533–550.
3. FROBENIUS, G., *Über die Charaktere der alternierenden Gruppe*, Sitzber. Preuss. Akad., 1901, pp. 303–315.
4. LITTLEWOOD, D. E., *The Theory of Group Characters*, Oxford, 1940.
5. MURNAGHAN, F. D., *The Theory of Group Representations*, Baltimore, 1938.
6. WEYL, H., *The Classical Groups*, Princeton, 1939.
7. YOUNG, A., *Quantitative substitutional analysis* I, Proceedings of the London Mathematical Society (1), vol. 33(1900), pp. 97–146; II, (1), vol. 34(1902), pp. 361–397; III, (2), vol. 28(1928), pp. 255–292; IV, (2), vol. 31(1930), pp. 253–272; V, (2), vol. 31(1930), pp. 273–288; VI, (2), vol. 34(1932), pp. 196–230; VII, (2), vol. 36(1934), pp. 304–368; VIII, (2), vol. 37(1934), pp. 441–495.

INSTITUTE FOR ADVANCED STUDY AND UNIVERSITY OF MICHIGAN.

PERIODIC TYPES OF TRANSFORMATIONS

By D. W. Hall and J. L. Kelley

1. Introduction. We consider in this paper a compact metric space X and a homeomorphism $T(X) \subset X$. For any subset Y of X we define the *orbit* of Y by the equation

$$O(Y) = \sum_{-\infty}^{\infty} T^i(Y).$$

The types of periodic transformations[1] which we shall consider are defined as follows:

(a) T is *pointwise periodic* provided that for every point x of X the set $O(x)$ is finite. If there exists an integer N such that no $O(x)$ contains more than N points, then T is said to be *periodic*.

(b) T is *almost periodic* provided that for every $\epsilon > 0$ there exists a positive integer n such that $\rho(x, T^n(x)) < \epsilon$ for every x in X.

(c) T is *pointwise almost periodic* provided that for each x in X and each $\epsilon > 0$ there exists a positive integer $n = n(x, \epsilon)$ such that $\rho(x, T^n(x)) < \epsilon$.

(d) T is *strongly pointwise almost periodic* provided that for any x in X and any $\epsilon > 0$ there exists an integer $K = K(x, \epsilon)$ such that every block of K successive integers contains an integer n for which $\rho(x, T^n(x)) < \epsilon$.

(e) T is *strongly almost periodic* provided that the integers K and n in (d) can be chosen independent of x. Besides these types of periodic transformations we shall deal with one type of non-periodic transformation defined as follows:

(f) T is said to be *regular*[2] provided that for any $\epsilon > 0$ there exists a $\delta > 0$ such that if $\rho(x, y) < \delta$ then $\rho(T^n(x), T^n(y)) < \epsilon$ for all integers n.

We note that the uniformity condition in *strongly almost periodic* is necessary in the sense that there exists a pointwise periodic and almost periodic transformation failing to satisfy this uniformity condition.

In the second section of this paper we investigate the conditions under which a collection of point orbits under a homeomorphism T will be continuous, i.e., each orbit is closed and if $\lim x_i = x$, then $\lim O(x_i) = O(x)$. We see at once from the condition that each $O(x)$ be closed that we must have a pointwise periodic transformation. Otherwise, it would follow that some $O(x)$ were infinite

Received May 15, 1941; presented to the American Mathematical Society, September 1939. This paper was started when the authors were in residence at the University of Virginia, the first named author as a National Research Fellow.

[1] See Deane Montgomery, *Pointwise periodic homeomorphisms*, American Journal of Mathematics, vol. 59(1937), pp. 118–120. See also W. L. Ayres, *On transformations having periodic properties*, Fundamenta Mathematicae, vol. 33(1939), pp. 95–103.

[2] See B. Kerékjártó, Acta Szeged, vol. 6(1932–34), p. 236.

and hence contained a limit point of itself. By applying the proper powers of T to this limit point it would follow that $O(x)$ was a perfect set. This is impossible since $O(x)$ is countable.

In the third section we study a continuous transformation $T(X) \subset X$. We say that a set Y is *irreducibly fixed* under T provided Y is closed, $T(Y) = Y$ and for any proper closed subset Z of Y, $T(Z) \neq Z$. We deduce necessary and sufficient conditions that a subset Y of X be irreducibly fixed under T and also necessary and sufficient conditions that a point x of X belong to an irreducibly fixed set under T. We get the interesting result that X contains a subset Y such that (i) Y is finite and T is periodic on Y, or (ii) Y is perfect and T is pointwise almost periodic on Y.

The final section characterizes completely the pointwise periodic transformations $F(M) = M$ defined on a plane locally connected continuum M not separated by any pair of its points.

2. Continuous collections of orbits.

We devote this section to the proof of the following theorem:

THEOREM 1. *If X is a compact metric space and $T(X) = X$ is pointwise periodic, then all of the following conditions are equivalent:*

(i) *the collection of point orbits under T is a continuous collection,*

(ii) *T is strongly almost periodic,*

(iii) *T is regular,*

(iv) *the orbit under T of every closed subset of X is closed.*

Proof. We first show that each of (i), (ii), and (iii) implies (iv). If condition (iv) fails, there exists a closed set $Y \subset X$ with $O(Y)$ not closed, and we can choose y, $y_i \in Y$, and integers n_i, with $\lim y_i = y$, $\lim T^{n_i}(y_i) = z \bar{\epsilon} O(Y)$. Since $O(y) \subset O(Y)$, $z \epsilon' O(y)$ and (i) fails. Let d be the least distance from a point of $O(z)$ to Y. If T is strongly almost periodic choose K_i, K, $0 \leqq K_i \leqq K$ such that $\rho(T^{n_i+K_i}(x), x) < d/2$ for all $x \in X$. Infinitely many of the K_i are equal, say to K_0, and $\lim T^{K_0}(T^{n_i}(y_i)) = T^{K_0}(z) \in O(z)$. But $\rho(T^{K_0+n_i}(y_i), y_i) < d/2$ for infinitely many i and hence $T^{K_0}(z)$ is at most $d/2$ from Y, which is a contradiction. Hence (ii) fails. That condition (iii) fails to hold is clear, for $\lim \rho(y_i, y) = 0$, but $\lim \rho(T^{n_i}(y_i), T^{n_i}(y)) \geqq d$.

It is trivial that (iv) implies (i) for we need only take for our closed set Y a convergent sequence of points of X together with its limit point. Hence (i) and (iv) are equivalent.

Suppose now that (i) holds. For an arbitrary x in X let $4d$ be the minimum distance between any two points of $O(x)$, and $p(x)$ the period of x under T. Let V be the spherical d-neighborhood of $O(x)$. By uniform continuity we can find a positive number d' less than d such that for any two points x and y of X such that $\rho(x, y) < d'$ we have $\rho(T^n(x), T^n(y)) < d$ for all n not exceeding $p(x)$. By (i) there exists a d'' less than d' such that any point y in the d''-neighborhood of x has its orbit in the d'-neighborhood of $O(x)$. It follows at once that if W

be the d''-neighborhood of x, and if W' be the sum of all images of W containing the point x then W' is an open set containing x, of diameter less than d', and $T^{p(x)}(W') = W'$. By the Heine-Borel Theorem we may cover X by a finite number of such neighborhoods $W_1', W_2', \cdots, W_n'$ such that for each i, $T^{p(x_i)}(W_i') = W_i'$. It follows at once that (iii) holds. We see that (ii) also holds since if K be any positive integer and p the product of the $p(x_i)$ $(i = 1, 2, \cdots, n)$, then T^{Kp} is within ϵ of the identity if each of the W_i' is taken of diameter less than ϵ. Thus (i) implies (ii) and (iii) and hence (i), (ii), (iii) and (iv) are equivalent.

Remark. Suppose T is a regular homeomorphism on a compact space. Define $\rho_n(x, y) = \rho(T^n(x), T^n(y))$. Clearly ρ_n is a metric equivalent to $\rho(x, y)$, i.e., $\rho_n(x_i, x) \to 0$ if and only if $\rho(x_i, x) \to 0$. Hence $\rho^*(x, y) = \operatorname{l.u.b.}_n \rho_n(x, y)$ satisfies the metric axioms, and by regularity is equivalent to $\rho(x, y)$, and $\rho^*(T^n(x), T^n(y)) = \rho^*(x, y)$ for all n. Any isometry is surely regular. Hence, regularity is equivalent to isometry and in particular *any pointwise periodic transformation satisfying one of* (i), (ii), (iii) *and* (iv) *is equivalent to a pointwise periodic isometry.*

We might also state this property of pointwise periodic transformations in a slightly different fashion. If any one of (i), (ii), (iii), (iv) hold the collection $\{T^n\}$ of transformations is equicontinuous, and hence the closure of this set in X^X is a compact transformation group G. The orbit under G of $x \in X$, i.e., the set of all images of x under transformations of G, is easily seen to be identical with $O(x)$. Hence *a pointwise periodic transformation T satisfies any one of* (i), (ii), (iii), (iv) *if and only if the collection $\{T^n\}$ is dense in a compact transformation group G, where the orbit of any $x \in X$ under G is identical with its orbit under T.*

3. On the existence of pointwise almost periodic transformations.[3]

The following Lemma has recently been established by one of us:[4]

LEMMA 1. *If X is compact and metric and $T(X) \subset X$ is a continuous transformation, then $\prod_1^\infty T^i(X)$ is fixed, and the property of being a fixed set under T is inducible.*

This Lemma yields at once the following characterization of irreducibly fixed subsets of X under T.

THEOREM 2. *In order that Y be irreducibly fixed under T it is necessary and sufficient that the images under T of every point of Y be dense in Y.*

Proof. The sufficiency is trivial. To prove the necessity assume that Y is

[3] It has been pointed out to the authors that the theorems of §3 are precisely analogous to certain results of G. D. Birkhoff for continuous flows. See *Dynamical Systems*, American Mathematical Society, Coll. Publications, vol. 9, pp. 195–205.

[4] See J. L. Kelley, Duke Mathematical Journal, vol. 5(1939), pp. 535–537.

irreducibly fixed under T, let x be an arbitrary point of Y and $L = \overline{\sum_0^\infty T^i(x)}$; then $T(L) \subset L$, and hence by Lemma 1, L contains a fixed set. Thus $L = Y$, and the proof is complete.

As an immediate consequence of Theorem 2 we get

THEOREM 3. *If X is compact and metric and $T(X) \subset X$ is continuous then there exists a subset Y of X such that either*

(i) *Y is finite and T is periodic on Y, or*

(ii) *Y is perfect and T is pointwise almost periodic on Y.*

THEOREM 4. *If X is compact and metric and $T(X) \subset X$ is continuous, then a necessary and sufficient condition that a given point x of X belong to some irreducibly fixed set under T is that T be strongly pointwise almost periodic at x.*

Proof. Let Y be irreducibly fixed under T and x an arbitrary point of Y. Suppose that T is not strongly pointwise almost periodic at x. Then for some positive ϵ we can find, for every integer K, an integer n_K such that $\sum_{i=1}^{K} T^{n_K+i}(x) \cdot V_\epsilon(x) = 0$. We may suppose $T^{n_K}(x)$ converges to a point y of Y as K becomes infinite. It follows easily that $O(y)$ is disjoint with $V_\epsilon(x)$, which contradicts Theorem 2 as Y is irreducibly fixed.

Conversely, suppose that T is strongly pointwise almost periodic at x and that $Y = \overline{\sum_0^\infty T^i(x)}$. If y is an arbitrary point of Y, then we must show that the images of y are dense in Y. This will evidently follow if we show that x is a limit point of the set of image points of y. To see this, for $\epsilon > 0$, pick a sequence of positive integers n_i such that $\lim T^{n_i}(x) = y$. Then since T is strongly pointwise almost periodic we can find an integer K, and then for each i an integer K_i, $1 \leqq K_i \leqq K$, such that $T^{n_i+K_i}(x) \in V_{\epsilon/2}(x)$. Now infinitely many of the K_i must be equal to a certain fixed K_0. Thus $\lim T^{K_0}(T^{n_i}(x)) = T^{K_0}(y) \in V_\epsilon(x)$. Hence x is a limit point of $\sum_0^\infty T^i(y)$ and the theorem is proved.

4. **On pointwise periodic transformations in the plane.** In this section we consider a locally connected continuum M which is embedded in the surface of a sphere S, and which has the property that no two points of M separate M. We shall classify completely all pointwise periodic transformations $T(M) = M$. This is most easily done by showing that there exists a pointwise periodic homeomorphism $Z(S) = S$ which is identical with T on M. It then follows at once from known results[5] that Z is periodic on S and hence T is periodic on M. But the action of a periodic transformation of a sphere S into itself has been com-

[5] See Deane Montgomery, loc. cit.

pletely and simply described,[6] and hence we can describe the action of Z (i.e., T) on M. We begin with the following Lemma.

LEMMA 2. *If $T(M) = M$ is pointwise periodic then T can be extended to the sphere S in such a way as to remain pointwise periodic.*

Proof. We note that the boundary of every complementary domain of M on S is a simple closed curve,[7] and it follows at once from our assumption that no two points of M separate M, that the boundaries of no two complementary domains of M on S can have more than one point in common. Thus, by a theorem of Adkisson,[8] there exists a homeomorphism $Z(S) = S$ which is identical with T on M.

Let J be the boundary of an arbitrary complementary domain of M on S and p and q arbitrary points of J. It follows that $T^m(J) = J_m$ is the boundary of a complementary domain[9] of M on S where m is the least common multiple of the periods of p and q under T. Since J and J_m have at least two points in common they must be identical. Hence $T^m(J) = J$ and it follows at once[10] that under T^m the set J either has exactly the two fixed points p and q or consists entirely of fixed points. If only p and q are fixed under T^m then every other point of J has period 2 under this transformation.

Let the orbit of J under T consist of the $k + 1$ simple closed curves $J_0, J_1, \cdots, J_k$, where $J_i = T^i(J)$ for each i and $J = T(J_k)$. For $i = 0, 1, 2, \cdots, k$ let E_i be the closure of the complementary domain of M on S bounded by J_i. Thus E_i is a 2-cell. Our proof will be complete if we show that the transformation T can be extended to the 2-cell E_0 while maintaining its pointwise periodicity. Suppose first that the J_i are all distinct. If all the points on J have the same period the extension is easy. Otherwise, there exist two points r and s on J having the same period while every other point on J has double this period, and the extension can again be made by elementary methods.

Finally, suppose that the J_i are not all distinct, and let m be the minimum value of the period function on J. It follows easily that there exists a unique point p on J having period m, since if two such points existed the J_i would be all distinct. Let n be the minimum value of the period function on the set $J - p$. If every point of $J - p$ has period n under T, then this transformation is easily extended to E_0. Otherwise, there exists a unique point q in $J - p$ having period n under T and every other point of $J - P$ has period $2n$ under T. The

[6] See (1) B. Kerékjártó, Mathematische Annalen, vol. 80(1919), pp. 36–38. (2) L. E. J. Brouwer, Mathematische Annalen, vol. 80(1919), pp. 39–41. (3) S. Eilenberg, Fundamenta Mathematicae, vol. 22(1934), pp. 28–41.

[7] See R. L. Moore, Fundamenta Mathematicae, vol. 6(1924), p. 212.

[8] See V. W. Adkisson, Comptes Rendus de la Société des Sciences de Varsovie, vol. 23 (1930), p. 168.

[9] See V. W. Adkisson, loc. cit., p. 167.

[10] See S. Eilenberg, loc. cit.

transformation is once more easily extended to E_0 and the proof of the lemma is complete.

As pointed out above this Lemma has the following immediate consequence:

THEOREM 5. *If $T(M) = M$ is pointwise periodic, where M is any locally connected continuum embedded in S such that no two points of M separate M, then T is periodic.*

We shall assume in the remainder of this section that the hypotheses of Theorem 5 are satisfied, and denote by K the set of all fixed points of M under T. By $Z(S) = S$ we denote the transformation set up in Lemma 2, and we signify by K' the set of all fixed points of S under Z. It has been shown[11] that one of the following cases must occur:

(a) $K' = S$.

(b) *K' is a simple closed curve on S, and Z is a reflection of S in this simple closed curve. Every point in $S - K'$ has period 2 under Z.*

(c) *K' consists of exactly two points a and b, which may be regarded as the poles of S. The transformation consists of a rotation of S around these poles and every point in $S - K'$ has the same period.*

(d) $K' = 0$. *In this case S has exactly two points a and b each of which is of period 2 under Z, and Z^2 is the transformation described in* (c).

Now K is a subset of K', hence we can read off numerous theorems regarding T from the results given above for Z. We give a single example in order to illustrate the type of result.

THEOREM 6. *If K is a proper subset of M containing more than two points, the following statements are true.*

(a) *K is a strong symmetric cut set[12] of M. In fact, $M - K$ has exactly two components R_1 and R_2.*

(b) *K is an irreducible cutting of M, so that $K = F(R_1) = F(R_2)$.*

(c) $T(R_1) = R_2$, $T(R_2) = R_1$.

(d) *There exists a simple closed curve J in S such that K is a subset of J.*

BROWN UNIVERSITY AND NOTRE DAME UNIVERSITY.

[11] See S. Eilenberg, loc. cit.

[12] See W. Dancer, Fundamenta Mathematicae, vol. 27(1936), pp. 123–135.

A PARTITION FUNCTION CONNECTED WITH THE MODULUS FIVE

By Joseph Lehner

1. The purpose of this paper is to derive a convergent series for $p_1(n)$ and $p_2(n)$, the number of partitions of a positive integer n into summands of the form $5l \pm 1$ and $5l \pm 2$, respectively. These partition functions occur in the following theorems of I. Schur [7],[1] which can be regarded as further cases of a result due to Euler:[2]

A. *The number of partitions of n into summands whose minimal difference is two is equal to $p_1(n)$;*

B. *The number of partitions of n into summands whose minimal difference is two and in which the summand one does not occur is equal to $p_2(n)$.*

It is convenient to treat $p_1(n)$ and $p_2(n)$ together. The other possible case for the modulus 5, that in which all summands are divisible by 5, reduces trivially to the *unrestricted* partition function $p(n)$: $p_0(n) = p(n/5)$. $p(n)$, on the other hand, has been fully discussed [2].

We follow the method of Rademacher [3]. Consider the generating functions

$$
\begin{aligned}
F_a(x) &= \prod_{m=0}^{\infty} (1 - x^{5m+a})^{-1} \prod_{m=1}^{\infty} (1 - x^{5m-a})^{-1} \\
&= 1 + \sum_{n=1}^{\infty} p_a(n) x^n \qquad (a = 1, 2),
\end{aligned} \tag{1.1}
$$

which converge inside the unit circle. In order to determine the asymptotic behavior of $F_a(x)$ near a "rational point" on a circle concentric to the unit circle but interior to it, we subject x to the transformation[3] $x \to x'$, where

$$
x = \exp\left[2\pi i \frac{h + iz}{k}\right], \qquad x' = \exp\left[2\pi i \frac{h' + iz^{-1}}{k}\right]. \tag{1.2}
$$

Here $\Re z > 0$, h and k are coprime integers satisfying $0 \leqq h < k$, and h' is any fixed solution of

$$
hh' \equiv -1 \pmod{k}. \tag{1.21}
$$

Received May 15, 1941; presented to the American Mathematical Society, May 2, 1941. The author wishes to acknowledge his deep indebtedness to Professor Hans Rademacher, who suggested this problem and gave much helpful advice connected with its solution.

[1] Numbers in square brackets refer to the bibliography at the end of this paper.

[2] Namely, the number of partitions of n into unequal parts (i.e., parts whose minimal difference is one) is equal to the number of partitions of n into odd parts (parts congruent ± 1 modulo 2).

[3] This amounts to a modular transformation of $F_a(x)$ considered as a function of τ: $x = \exp(2\pi i \tau)$. See §7.

We then derive in §§2–4 a functional equation (4.7) connecting $F_a(x)$ and $F_b(x')$ for all k divisible by 5 (b equals 1 or 2 depending on h). If $5 \nmid k$, we write

$$(1.3) \qquad x'' = \exp\left(2\pi i \frac{H' + iz^{-1}}{K}\right), \qquad HH' \equiv -1 \pmod{k},$$

and find a similar relation (6.7) between $F_a(x)$ and certain new functions $H_a(x'')$.

The transformation equations (4.7) and (6.7) contain certain complicated roots of unity, $\omega_a(h, k)$ and $\chi_a(h, k)$, respectively, whose complete definition is given in (4.71) and (6.71). The application of Rademacher's method requires that certain sums of these roots of unity, for example

$$\sum_h \omega_a(h, k) \exp(-2\pi i hnk^{-1}),$$

$$5 \mid k;\, h \bmod k,\ (h, k) = 1,\ h \equiv d \pmod{5}$$

be estimated more sharply than by the trivial $O(k)$. This is done in Part II by reducing the sums in question to incomplete Kloosterman sums, for which suitable estimates have been provided by Estermann [1] and Salié [6].

After these preparations we derive in Part III a convergent series and an asymptotic formula for $p_a(n)$.

I. The transformation equations

2. Our first object is to derive a transformation equation for $F_a(x)$ in (1.1). Now $F_a(x)$ differs by only elementary factors from a certain modular function,[4]

$$(2.1) \qquad -i \exp\left[\pi i\tau \frac{6a - 5}{6}\right] F_a(x) = \frac{\eta(5\tau)}{\vartheta_1(a\tau \mid 5\tau)}, \qquad x = \exp(2\pi i\tau),$$

where $\eta(\tau)$ is the elliptic modular function of Dedekind,

$$(2.11) \qquad \eta(\tau) = \exp\left(\frac{\pi i\tau}{12}\right) \prod_{m=1}^{\infty} (1 - \exp(2\pi i m\tau)).$$

Doubtless the desired transformation can be obtained by using known results of ϑ-function theory. It is more convenient, however, to use an independent method developed by Rademacher [7] (§1).

We take x as in (1.2) and consider the case $5 \mid k$ first. $F_a(x)$ is regular and zero free inside the unit circle, hence $\log F_a(x)$ is single-valued in the same region if we first choose a specific branch of the logarithm, say the one given by $\log F(0) = 0$. Then

$$(2.21) \qquad \begin{aligned} G_a(x) = \log F_a(x) &= -\sum_{m=0}^{\infty} \log(1 - x^{5m+a}) - \sum_{m=1}^{\infty} \log(1 - x^{5m-a}) \\ &= \sum_{n=1}^{\infty}\sum_{m=0}^{\infty} \frac{x^{(5m+a)n}}{n} + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{x^{(5m-a)n}}{n}. \end{aligned}$$

[4] For definitions of the ϑ-functions, see [8], p. 234 ff. What is here referred to as $\eta(\tau)$ is there denoted by $h(\tau)$.

Setting

$$5m \pm a = qk + \mu_a, \qquad n = rk + \nu,$$

(2.22) $\mu_a \equiv \pm a \pmod 5$, $0 < \mu_a < k$; $\nu = 1, 2, \cdots, k$; $q, r = 0, 1, 2, \cdots,$

we have from (2.21)

$$(2.23) \qquad G_a(x) = \sum_{\mu_a, \nu} \exp\left(2\pi i \frac{h}{k} \mu\nu\right) \sum_{q,r} \frac{1}{rk + \nu} \exp\left\{-\frac{2\pi z}{k}(qk + \mu)(rk + \nu)\right\},$$

where μ_a, ν, q and r run over the ranges described above. For simplicity of notation we have written μ for μ_a in the summands themselves.

Application of Mellin's formula to (2.23) gives

$$\begin{aligned} G_a(x) &= \sum_{\mu_a, \nu} \exp\left(2\pi i \frac{h}{k} \mu\nu\right) \sum_{q,r} \frac{1}{rk + \nu} \frac{1}{2\pi i} \int_{(\frac{3}{2})} \frac{\Gamma(s) k^s \, ds}{(2\pi z)^s (qk + \mu)^s (rk + \nu)^s} \\ (2.24) \qquad &= \frac{1}{2\pi i} \sum_{\mu_a, \nu} \exp\left(2\pi i \frac{h}{k} \mu\nu\right) \int_{(\frac{3}{2})} \frac{\Gamma(s)}{(2\pi z)^s k^{s+1}} \sum_{q=0}^{\infty} \left(q + \frac{\mu}{k}\right)^{-s} \sum_{r=0}^{\infty} \left(r + \frac{\nu}{k}\right)^{-s-1} ds \\ &= \frac{1}{2\pi i k} \sum_{\mu_a, \nu} \exp\left(2\pi i \frac{h}{k} \mu\nu\right) \int_{(\frac{3}{2})} \frac{\Gamma(s)}{(2\pi z k)^s} \zeta\left(s, \frac{\mu}{k}\right) \zeta\left(s + 1, \frac{\nu}{k}\right) ds, \end{aligned}$$

where $\zeta(s, w)$, $0 < w \leqq 1$, is the Hurwitz zeta-function. By z^s we mean exp $(s \log z)$, and since $\Re z > 0$ we can take $|\, \Im \log z \,| < \frac{1}{2}\pi$. Moreover, the symbol (u) under an integral sign means that the range of integration is from $u - i\infty$ to $u + i\infty$.

In the transformation equation of $\zeta(s, w)$, we put $w = \mu/k$ and obtain

$$\begin{aligned} \zeta\left(s, \frac{\mu}{k}\right) = \Gamma(1 - s) \frac{2}{(2\pi k)^{1-s}} \Bigg(\sin \frac{\pi s}{2} \sum_{\lambda=1}^{k} \cos 2\pi \frac{\lambda\mu}{k} \cdot \zeta\left(1 - s, \frac{\lambda}{k}\right) \\ (2.31) \qquad + \cos \frac{\pi s}{2} \sum_{\lambda=1}^{k} \sin 2\pi \frac{\lambda\mu}{k} \cdot \zeta\left(1 - s, \frac{\lambda}{k}\right)\Bigg). \end{aligned}$$

Moreover, we note that

$$(2.32) \qquad \sum_{\mu_a} \exp\left(2\pi i \frac{h}{k} \mu\nu\right) \cos 2\pi \frac{\lambda\mu}{k} = \sum_{\mu_a} \cos 2\pi \frac{h}{k} \mu\nu \cos 2\pi \frac{\lambda\mu}{k},$$

$$(2.33) \qquad \sum_{\mu_a} \exp\left(2\pi i \frac{h}{k} \mu\nu\right) \sin 2\pi \frac{\lambda\mu}{k} = i \sum \sin 2\pi \frac{h}{k} \mu\nu \sin 2\pi \frac{\lambda\mu}{k}.$$

If we apply (2.31)–(2.33) to (2.24) we obtain after simplification

$$\begin{aligned} G_a(x) = \frac{1}{4\pi i k^2} \sum_{\mu_a, \nu, \lambda} \cos 2\pi \frac{h}{k} \mu\nu \cos 2\pi \frac{\lambda\mu}{k} \int_{(\frac{3}{2})} \frac{\zeta(1 - s, \lambda/k) \zeta(1 + s, \nu/k)}{z^s \cos \frac{1}{2}\pi s} ds \\ + \frac{1}{4\pi k^2} \sum_{\mu_a, \nu, \lambda} \sin 2\pi \frac{h}{k} \mu\nu \sin 2\pi \frac{\lambda\mu}{k} \int_{(\frac{3}{2})} \frac{\zeta(1 - s, \lambda/k) \zeta(1 + s, \nu/k)}{z^s \sin \frac{1}{2}\pi s} ds; \end{aligned}$$

λ takes the values $1, 2, \cdots, k$.

In the sums of (2.34) we now introduce the new summation letter μ_b by

$$\text{(2.41)} \qquad \begin{cases} \mu_a \equiv h'\mu_b \pmod{k}, \\ 0 < \mu_b < k, \end{cases}$$

where h' is given by (1.21). If $h \equiv \pm 1 \pmod 5$ then $h' \equiv \mp 1 \pmod 5$ (since $5 \mid k$) and μ_b runs over the same system of values as μ_a in some order. That is, $b = a$ in this case. But if $h \equiv \pm 2 \pmod 5$ then $h' \equiv \pm 2 \pmod 5$ and μ_b runs over the system "complementary" to μ_a. More precisely, let (a, a^*) be either the pair (1, 2) or (2, 1). Then

$$\text{(2.42)} \qquad b = b(h) = \begin{cases} a, & h \equiv \pm 1 \pmod 5, \\ a^*, & h \equiv \pm 2 \pmod 5. \end{cases}$$

In addition we replace s by $-s$ and get

$$\text{(2.5)} \qquad \begin{aligned} G_a(x) = \frac{1}{4\pi i k^2} \sum_{\mu_b,\nu,\lambda} \cos 2\pi \frac{\mu\nu}{k} \cos 2\pi \frac{h'\lambda\mu}{k} \int_{(-\frac{3}{2})} \frac{\zeta(1+s, \lambda/k)\zeta(1-s, \nu/k)}{z^{-s} \cos \frac{1}{2}\pi s}\, ds \\ + \frac{1}{4\pi k^2} \sum_{\mu_b,\nu,\lambda} \sin 2\pi \frac{\mu\nu}{k} \sin 2\pi \frac{h'\lambda\mu}{k} \int_{(-\frac{3}{2})} \frac{\zeta(1+s, \lambda/k)\zeta(1-s, \nu/k)}{z^{-s} \cos \frac{1}{2}\pi s}\, ds. \end{aligned}$$

If we now shift the path of integration back to $\sigma = \Re s = \frac{3}{2}$, the sums take the same form as in (2.34) except that μ_a is replaced by μ_b, h by h', λ and ν are interchanged, and z becomes $1/z$. The displacement of the path of integration is easily justified by the presence of $\cos \frac{1}{2}\pi s$ and $\sin \frac{1}{2}\pi s$ in the denominators of the integrands. Hence, recalling the definition (1.2) of x', we obtain

$$\text{(2.6)} \qquad \begin{aligned} G_a(x) &= G_b(x') - 2\pi i (R_1 + R_2), \\ R_1 &= \sum_{-\frac{3}{2}<\sigma<\frac{3}{2}} \operatorname{Res} I_1, \qquad R_2 = \sum_{-\frac{3}{2}<\sigma<\frac{3}{2}} \operatorname{Res} I_2, \end{aligned}$$

I_1 and I_2 being the integrands obtained by removing the integral sign to the extreme left in the first and second terms, respectively, of the right member of (2.5).

3. Recalling that

$$\zeta(s, w) = \frac{1}{s-1} + \gamma(w) + O(s-1), \qquad \gamma(w) = -\frac{\Gamma'(w)}{\Gamma(w)},$$

we can obtain from (2.5) and (2.6) an expression for R_1 by evaluating the residue of I_1 at its simple poles $s = \pm 1$ and at the double pole $s = 0$.

$$\text{(3.1)} \qquad \begin{aligned} R_1 = \frac{1}{4\pi i k^2} \sum_{\mu_b,\nu,\lambda} \cos 2\pi \frac{\mu\nu}{k} \cos 2\pi \frac{h'\lambda\mu}{k} \Big\{ \Big(-\frac{2z}{\pi} \zeta\Big(0, \frac{\nu}{k}\Big) \zeta\Big(2, \frac{\lambda}{k}\Big) \Big) \\ + \Big(\frac{2}{\pi z} \zeta\Big(2, \frac{\nu}{k}\Big) \zeta\Big(0, \frac{\lambda}{k}\Big) \Big) + \Big(\gamma\Big(\frac{\nu}{k}\Big) - \gamma\Big(\frac{\lambda}{k}\Big) \Big) + (-\log z) \Big\} \\ = R_{11} + R_{12} + R_{13} + R_{14}. \end{aligned}$$

By letting $s \to 1$ in the transformation equation (2.31) we find

$$\sum_{\nu=1}^{k} \cos 2\pi \frac{\mu\nu}{k} \cdot \zeta\left(0, \frac{\nu}{k}\right) \tag{3.2}$$
$$= \frac{1}{2} \lim_{s \to 1} \frac{\zeta(s, \mu/k)}{\Gamma(1-s)} = \frac{1}{2} \lim_{s\to 1} \left\{\left(\frac{1}{s-1} + \cdots\right) \frac{(1-s)}{\Gamma(2-s)}\right\} = -\frac{1}{2}.$$

We denote by $\{a, b\}$ the unique real number defined by

$$\{a, b\} \equiv a \pmod{b}, \qquad 0 < \{a, b\} \leqq b. \tag{3.3}$$

Then letting $s \to -1$ in (2.31) we get

$$\sum_{\lambda=1}^{k} \cos 2\pi \frac{h'\lambda\mu}{k} \cdot \zeta\left(2, \frac{\lambda}{k}\right) = \sum_{\lambda=1}^{k} \cos 2\pi \frac{\lambda}{k} \cdot \{h'\mu, k\} \cdot \zeta\left(2, \frac{\lambda}{k}\right)$$
$$= -2\pi^2 k^2 \zeta\left(-1, \frac{\{h'\mu, k\}}{k}\right) \tag{3.41}$$
$$= 2\pi^2 k^2 \left(\frac{1}{2k^2} \{h'\mu, k\}^2 - \frac{1}{2k} \{h'\mu, k\} + \frac{1}{12}\right).$$

From (2.41) it follows that

$$\sum_{\mu_b} \{h'\mu, k\}^2 = \sum_{\mu_a} \{\mu, k\}^2 = \sum_{\mu_a} \mu^2 = \sum_{l=0}^{k/5-1} (5l+a)^2 + \sum_{l=1}^{k/5} (5l-a)^2$$
$$= \frac{k}{15}(2k^2 + 6a^2 - 30a + 25) = \frac{k}{15}(2k^2 + A), \tag{3.43}$$
$$A = 6a^2 - 30a + 25;$$

$$\sum_{\mu_b} \{h'\mu, k\} = \sum_{\mu_a} \{\mu, k\} = \sum_{\mu_a} \mu = \frac{k^2}{5}. \tag{3.44}$$

Thus we obtain from (3.1), (3.2), (3.41), (3.43), (3.44),

$$R_{11} = \frac{z}{4i}\left(\frac{1}{15k}(2k^2 + A) - \frac{k}{5} + \frac{1}{6} \cdot \frac{2k}{5}\right),$$
$$R_{11} = \frac{Az}{60ik}. \tag{3.5}$$

The value of R_{12} is easily found by using the above formulas:

$$R_{12} = -\frac{B}{60ikz}, \qquad B = 6b^2 - 30b + 2.5. \tag{3.6}$$

Moreover,

$$R_{13} = \frac{1}{4\pi i k^2} \sum_{\mu_b} \sum_{\nu=1}^{k} \cos 2\pi \frac{\mu\nu}{k} \cdot \gamma\left(\frac{\nu}{k}\right) \sum_{\lambda=1}^{k} \cos 2\pi \frac{h'\lambda\mu}{k} \tag{3.7}$$
$$- \frac{1}{4\pi i k^2} \sum_{\mu_b} \sum_{\lambda=1}^{k} \cos 2\pi \frac{h'\lambda\mu}{k} \cdot \gamma\left(\frac{\lambda}{k}\right) \sum_{\nu=1}^{k} \cos 2\pi \frac{\mu\nu}{k} = 0,$$

since the innermost sum of each term of the right member vanishes by virtue of $0 < \mu < k$, $(h, k) = (h', k) = 1$.

In the same way,

$$R_{14} = 0. \tag{3.8}$$

4. We shall now discuss R_2. From (2.6)

$$\begin{aligned} I_2 = \frac{1}{4\pi k^2} \frac{z^s}{\sin \frac{1}{2}\pi s} \sum_{\mu_h} \sum_{\nu=1}^{k} \sin 2\pi \frac{\mu\nu}{k} \\ \cdot \zeta\left(1 - s, \frac{\nu}{k}\right) \sum_{\lambda=1}^{k} \sin 2\pi \frac{h'\lambda\mu}{k} \cdot \zeta\left(1 + s, \frac{\lambda}{k}\right). \end{aligned} \tag{4.1}$$

The only pole of I_2 in the rectangle $-\frac{3}{2} < \sigma < \frac{3}{2}$ is at $s = 0$. This pole is apparently of the third order, but, as we shall show immediately, the inner sums of (4.1) are actually regular at $s = 0$ so that we have in reality only a simple pole there. For this purpose we have recourse to the transformation equation (2.31) once more.

$$\begin{aligned} \sum_{\nu=1}^{k} \sin 2\pi \frac{\mu\nu}{k} \cdot \zeta\left(1 - s, \frac{\nu}{k}\right) &= \frac{(2\pi k)^{1-s}}{2\Gamma(1 - s) \cos \frac{1}{2}\pi s} \zeta\left(s, \frac{\mu}{k}\right) \\ &\quad - \frac{1}{\cos \frac{1}{2}\pi s} \sum_{\nu=1}^{k} \cos 2\pi \frac{\mu\nu}{k} \sin \frac{\pi s}{2} \cdot \zeta\left(1 - s, \frac{\nu}{k}\right), \end{aligned}$$

$$\begin{aligned} &\lim_{s \to 0} \sum_{\nu=1}^{k} \sin 2\pi \frac{\mu\nu}{k} \cdot \zeta\left(1 - s, \frac{\nu}{k}\right) \\ &= \pi k \zeta\left(0, \frac{\mu}{k}\right) - \sum_{\nu=1}^{k} \cos 2\pi \frac{\mu\nu}{k} \lim_{s \to 0} \left\{\left(\frac{\pi s}{2} - \cdots\right)\left(-\frac{1}{s} + \cdots\right)\right\} \\ &= \pi k \left(\frac{1}{2} - \frac{\mu}{k}\right) + \frac{\pi}{2} \sum_{\nu=1}^{k} \cos 2\pi \frac{\mu\nu}{k} = \pi k \left(\frac{1}{2} - \frac{\mu}{k}\right) + \frac{\pi k}{2} \delta\left(\frac{\mu}{k}\right), \end{aligned} \tag{4.21}$$

where

$$\delta(x) = \begin{cases} 1, & \text{if } x \text{ is an integer,} \\ 0, & \text{otherwise.} \end{cases} \tag{4.22}$$

Since $0 < \mu < k$, we have $\delta(\mu/k) = 0$, and therefore

$$\sum_{\nu=1}^{k} \sin 2\pi \frac{\mu\nu}{k} \cdot \zeta\left(1 - s, \frac{\nu}{k}\right) = \pi k \left(\frac{1}{2} - \frac{\mu}{k}\right) + O(s). \tag{4.23}$$

The sum over λ in (4.1) can be treated in the same way.

$$\begin{aligned} \lim_{s \to 0} \sum_{\lambda=1}^{k} \sin 2\pi \frac{h'\lambda\mu}{k} \cdot \zeta\left(1 + s, \frac{\lambda}{k}\right) &= \lim_{s \to 0} \sum_{\lambda=1}^{k} \sin 2\pi \frac{\lambda}{k} \{h'\mu, k\} \cdot \zeta\left(1 + s, \frac{\lambda}{k}\right) \\ &= \pi k \left(\frac{1}{2} - \frac{\{h'\mu, k\}}{k}\right) + \frac{\pi k}{2} \delta\left(\frac{\{h'\mu, k\}}{k}\right). \end{aligned} \tag{4.31}$$

From the definition (3.3) of $\{h'\mu, k\}$ we have

$$\frac{\{h'\mu, k\}}{k} = \frac{h'\mu}{k} - \left[\frac{h'\mu}{k}\right],$$

$$\delta\left(\frac{\{h'\mu, k\}}{k}\right) = 0;$$

hence,

$$\sum_{\lambda=1}^{k} \sin 2\pi \frac{h'\lambda\mu}{k} \cdot \zeta\left(1 + s, \frac{\lambda}{k}\right) = \pi k\left(\frac{1}{2} - \frac{h'\mu}{k} + \left[\frac{h'\mu}{k}\right]\right) + O(s). \tag{4.32}$$

The sums on ν and λ in (4.1) are therefore regular at $s = 0$ and we have indeed a simple pole there. It is now an easy matter to calculate the residue.

$$R_2 = \frac{1}{2}\sum_{\mu_b}\left(\frac{1}{2} - \frac{\mu}{k}\right)\left(\frac{1}{2} - \frac{h'\mu}{k} + \left[\frac{h'\mu}{k}\right]\right). \tag{4.4}$$

We define[5] for real x,

$$((x)) = \begin{cases} x - [x] - \frac{1}{2}, & \text{for } x \text{ not an integer,} \\ 0, & \text{for integral } x, \end{cases} \tag{4.51}$$

that is,

$$((x)) = x - [x] - \tfrac{1}{2} + \tfrac{1}{2}\delta(x). \tag{4.52}$$

Clearly

$$((-x)) = -((x)), \tag{4.53}$$

$$((x + 1)) = ((x)). \tag{4.54}$$

Introducing (4.51) into (4.4), we get

$$R_2 = \frac{1}{2}\sum_{\mu_b}\left(\left(\frac{\mu}{k}\right)\right)\left(\left(\frac{h'\mu}{k}\right)\right) = \frac{1}{2}\sum_{\mu_b}\left(\left(-\frac{h\mu}{k}\right)\right)\left(\left(-\frac{hh'\mu}{k}\right)\right),$$

$$R_2 = -\frac{1}{2}\sum_{\mu_a}\left(\left(\frac{\mu}{k}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right), \tag{4.6}$$

where we have made use of $\mu_b \equiv -h\mu_a \pmod k$ from (2.41) and of (4.53), (4.54).

From (2.6), (3.1), (3.5), (3.6), (3.7), (3.8), and (4.6) we obtain after exponentiation the desired transformation equation for $F_a(x)$:

$$F_a(x) = \omega_a(h, k) \exp\left\{\frac{\pi}{30k}\left(\frac{B}{z} - Az\right)\right\} F_b(x'), \qquad 5 \mid k. \tag{4.7}$$

In (4.7), b is defined by (2.42), A by (3.43), B by (3.6), x and x' by (1.2), and

$$\omega_a(h, k) = \exp(\pi i r_a(h, k)) \tag{4.71}$$

[5] Cf. [5], formula (1.1).

with

$$r_a(h, k) = \sum_{\mu_a} \left(\left(\frac{\mu}{k}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right). \tag{4.72}$$

5. We must now take up the case $5 \nmid k$. This case is more difficult than the preceding one since $F_a(x)$ does not go over into $F_b(x')$ under transformation. However, we can determine explicitly the new function into which it is transformed. Since we use the same method as before, we shall abbreviate the discussion somewhat.

Here we choose x, x'', H, H', K as in (1.2), (1.3) and set[6]

$$\begin{gathered} 5m \pm a = qK + \mu_a, \qquad n = rk + \nu, \\ \mu_a \equiv \pm a \pmod 5, \qquad 0 < \mu_a < K; \\ \nu = 1, 2, \cdots, k; \qquad q, r = 0, 1, 2, \cdots. \end{gathered} \tag{5.1}$$

Then, as in §2,

$$\begin{aligned} G_a(x) = {} & \frac{1}{4\pi i k K} \sum_{\mu_a, \nu, \lambda} \cos 2\pi \frac{h}{k} \mu\nu \cos 2\pi \frac{\lambda\mu}{K} \int_{(\frac{3}{2})} \frac{\zeta(1 - s, \lambda/K)\zeta(1 + s, \nu/k)}{z^s \cos \frac{1}{2}\pi s} ds \\ & + \frac{1}{4\pi k K} \sum_{\mu_a, \nu, \lambda} \sin 2\pi \frac{h}{k} \mu\nu \sin 2\pi \frac{\lambda\mu}{K} \int_{(\frac{3}{2})} \frac{\zeta(1 - s, \lambda/K)\zeta(1 + s, \nu/k)}{z^s \sin \frac{1}{2}\pi s} ds, \end{aligned} \tag{5.2}$$

where λ takes the values $1, 2, \cdots, K$.

Now let μ^* be defined by

$$\begin{cases} \mu_a \equiv 5H'\mu^* \pmod k, \\ 0 < \mu^* \leqq k; \end{cases} \tag{5.3}$$

introducing μ^* into (5.2) and replacing s by $-s$ we find

$$\begin{aligned} G_a(x) = {} & \frac{1}{4\pi i k K} \sum_{\mu_a, \nu, \lambda} \cos 2\pi \frac{\mu^*\nu}{k} \cos 2\pi \frac{\lambda\mu}{K} \int_{(-\frac{3}{2})} \frac{\zeta(1 + s, \lambda/K)\zeta(1 - s, \nu/k)}{z^{-s} \cos \frac{1}{2}\pi s} ds \\ & + \frac{1}{4\pi k K} \sum_{\mu_a, \nu, \lambda} \sin 2\pi \frac{\mu^*\nu}{k} \sin 2\pi \frac{\lambda\mu}{K} \int_{(-\frac{3}{2})} \frac{\zeta(1 + s, \lambda/K)\zeta(1 - s, \nu/k)}{z^{-s} \sin \frac{1}{2}\pi s} ds. \end{aligned} \tag{5.41}$$

Cauchy's theorem then gives

$$G_a(x) = J_\alpha(x'') - 2\pi i(R_1 + R_2), \tag{5.42}$$

where $J_\alpha(x'')$ denotes the same sum as $G_a(x)$ except that the integrals are taken over $\sigma = \frac{3}{2}$. $\alpha = \alpha(a)$ will be defined later.

It is not obvious that the expression which we have denoted by $J_\alpha(x'')$ is a one-valued analytic function of $x'' = \exp\{2\pi i(H'k^{-1} + iK^{-1}z^{-1})\}$. From its

[6] Note that μ_a is defined as before, (2.22), except that its *modulus* is K instead of k. The modulus will always be indicated ($\mu_a = \mu_a(K)$) when necessary to avoid confusion. Throughout §§5–6 the modulus K is meant.

definition follows only the analyticity of $J_a(x'')$ in z for $\Re z > 0$. However, we shall now evaluate $J_a(x'')$ in a closed form from which the desired properties will follow at once.

To accomplish this we simply retrace our steps. In the expression for $J_a(x'')$ we replace $\cos(2\pi\lambda\mu/K)$ by $\exp(2\pi i\lambda\mu/K)$ and $\sin(2\pi\lambda\mu/K)$ by $i^{-1}\exp(2\pi i\lambda\mu/K)$. This is permissible in view of the reasoning of (2.32), (2.33). We then apply the transformation equation (2.31) for $\zeta(s, \mu^*/k)$ and obtain easily

$$J_a(x'') = \frac{1}{2\pi i K}\sum_{\mu_a,\lambda}\exp\left(2\pi i\frac{\lambda\mu}{K}\right)\int_{(\frac{3}{2})}\frac{z^s\Gamma(s)}{(2\pi k)^s}\zeta\left(s, \frac{\mu^*}{k}\right)\zeta\left(s+1, \frac{\lambda}{K}\right)ds,$$

and from Mellin's formula

$$J_a(x'') = \sum_{\mu_a,\lambda}\exp\left(2\pi i\frac{\lambda\mu}{K}\right)\sum_{q,r=0}^{\infty}\frac{1}{qK+\lambda}\exp\left[-\frac{2\pi}{Kz}(rk+\mu^*)(qK+\lambda)\right]. \tag{5.5}$$

Now the definition (5.1) shows that μ_a runs over a complete residue system modulo k twice in some order, since $5 \nmid k$; by (5.3) this must also be true of μ^*. We therefore put

$$\alpha k \equiv a \pmod 5, \qquad 0 < \alpha < 5, \tag{5.6}$$

and this, together with (5.3), yields

$$\mu_a \equiv 5H'\mu^* \pm \alpha k \pmod K,$$

the $\pm$ agreeing with $\mu_a \equiv \pm a \pmod 5$. The root of unity in (5.5) is then

$$\exp\left[2\pi i\left(\pm\frac{\lambda\alpha}{5} + \frac{\lambda H'\mu^*}{k}\right)\right]$$
$$= \exp\left[\pm\frac{2\pi i\alpha}{5}(qK+\lambda)\right]\exp\left[2\pi i\frac{H'}{k}(qK+\lambda)(rk+\mu^*)\right].$$

Hence, setting

$$qK+\lambda = m, \qquad rk+\mu^* = n, \qquad \rho = \rho(\alpha) = \exp\left(2\pi i\frac{\alpha}{5}\right), \tag{5.61}$$

we obtain

$$\begin{aligned} J_a(x'') &= \sum_{m,n=1}^{\infty}\frac{\rho^m x''^{mn}}{m} + \sum_{m,n=1}^{\infty}\frac{\bar\rho^m x''^{mn}}{m} \\ &= \log\prod_{n=1}^{\infty}(1-\rho x''^n)^{-1}\prod_{n=1}^{\infty}(1-\bar\rho x''^n)^{-1} = \log H_\alpha(x''), \end{aligned} \tag{5.7}$$

where

$$H_\alpha(x) = \prod_1^{\infty}(1-\rho x^n)^{-1}\prod_1^{\infty}(1-\bar\rho x^n)^{-1} = 1 + \sum_{\nu=1}^{\infty}c_\alpha(\nu)x^\nu. \tag{5.8}$$

(5.7) shows that the notation $J_\alpha(x'')$ of (5.42) is justified. $H_\alpha(x'')$ is regular at $x'' = 0$. Which of the functions $H_\alpha(x'')$ the function $F_\alpha(x)$ is transformed into depends partly on x itself but only through the residue of k modulo 5.

6. Next we discuss the residue sums, R_1 and R_2 of (5.42). From (5.41), (5.42) we deduce

$$R_1 = \frac{1}{4\pi ikK} \sum_{\mu_\alpha,\nu,\lambda} \cos 2\pi \frac{\mu^*\nu}{k} \cos 2\pi \frac{\lambda\mu}{K} \left\{ \left(-\frac{2z}{\pi} \zeta\left(0, \frac{\nu}{k}\right) \zeta\left(2, \frac{\lambda}{K}\right) \right) \right.$$

$$\left. + \left(\frac{2}{\pi z} \zeta\left(2, \frac{\nu}{k}\right) \zeta\left(0, \frac{\lambda}{K}\right) \right) + \left(\gamma\left(\frac{\nu}{k}\right) - \gamma\left(\frac{\lambda}{K}\right) \right) + (-\log z) \right\} \tag{6.1}$$

$$= R_{11} + R_{12} + R_{13} + R_{14}.$$

By a straightforward calculation similar to that of §3, using the remark following (5.5) and $0 < \mu < K$, we find

$$R_{11} = \frac{Az}{60ik}, \qquad R_{12} = -\frac{1}{60ikz}, \qquad R_{14} = 0, \tag{6.2}$$

where A is defined by (3.43).

There remains R_{13}. This is the difference of two triple sums of which the first vanishes, as is easily seen. From (2.31) we have

$$\sum_{\lambda=1}^{K} \cos 2\pi \frac{\lambda\mu}{K} \cdot \zeta\left(1 - s, \frac{\lambda}{K}\right)$$

$$= \frac{1}{\sin \frac{1}{2}\pi s} \frac{(2\pi K)^{1-s}}{2\Gamma(1-s)} \zeta\left(s, \frac{\mu}{K}\right) - \cot \frac{\pi s}{2} \sum_{\lambda=1}^{K} \sin 2\pi \frac{\lambda\mu}{K} \cdot \zeta\left(1 - s, \frac{\lambda}{K}\right).$$

Writing the Laurent expansion of each member around $s = 0$ and equating constant terms yield

$$\sum_{\lambda=1}^{K} \cos 2\pi \frac{\lambda\mu}{K} \cdot \gamma\left(\frac{\lambda}{K}\right) = 2K\left\{ -\log(2\pi K) \cdot \left(\frac{1}{2} - \frac{\mu}{K}\right) + A_1\left(\frac{\mu}{K}\right) \right\}$$

$$- \frac{2}{\pi} \left(B_1\left(\frac{\lambda}{K}\right) + \frac{\pi^2}{12} \right) \sum_{\lambda=1}^{K} \sin 2\pi \frac{\lambda\mu}{K}$$

$$= -2K\left(\frac{1}{2} - \frac{\mu}{K}\right)(\log(2\pi K) + \gamma) + 2A_1\left(\frac{\mu}{K}\right) \cdot K,$$

by virtue of $0 < \mu < K$. Here A_1 and B_1 are explained by

$$\zeta(s, w) = \tfrac{1}{2} - w + A_1(w) \cdot s + \cdots,$$

$$\zeta(1 - s, w) = -\frac{1}{s} + \gamma(w) + B_1(w) \cdot s + \cdots,$$

and

$$\gamma = \gamma(1) = -\frac{\Gamma'(1)}{\Gamma(1)}$$

is Euler's constant. The value of $A_1(w)$ is known:

$$A_1(w) = \zeta'(0, w) = \log \Gamma(w) - \tfrac{1}{2} \log (2\pi).$$

Furthermore

$$\sum_{\nu=1}^{k} \cos 2\pi \frac{\mu^*\nu}{k} = k\delta\left(\frac{\mu^*}{k}\right) = k\delta\left(\frac{\mu}{k}\right).$$

There are just two values of μ for which $\delta(\mu/k) \neq 0$, namely, $\mu_1 = \alpha k$, $\mu_2 = K - \alpha k$.

$$\mu_1 + \mu_2 = K,$$

$$\left(\frac{1}{2} - \frac{\mu_1}{K}\right) + \left(\frac{1}{2} - \frac{\mu_2}{K}\right) = 1 - \frac{\mu_1 + \mu_2}{K} = 0.$$

Using these results in (6.1) we find easily

$$R_{13} = -\frac{1}{4\pi ikK} \cdot k \cdot 2K(\log \Gamma(\mu_1/K) + \log \Gamma(\mu_2/K) - \log (2\pi)),$$

(6.3)

$$R_{13} = \frac{1}{2\pi i} \log \left(2 \sin \frac{\pi\alpha}{5}\right).$$

The residue sum R_2 is handled in the same manner as the corresponding sum in §4.

$$I_2 = \frac{-1}{4\pi kK} \cdot \frac{z^s}{\sin \frac{1}{2}\pi s} \sum_{\mu_a} \sum_{\nu=1}^{k} \sin 2\pi \frac{\mu^*\nu}{k}$$

(6.4)

$$\cdot\, \zeta\left(1 - s, \frac{\nu}{k}\right) \sum_{\lambda=1}^{K} \sin \frac{2\pi\lambda\mu}{K}\, \zeta\left(1 + s, \frac{\lambda}{K}\right).$$

Now

$$\sum_{\lambda=1}^{K} \sin 2\pi \frac{\lambda\mu}{K} \cdot \zeta\left(1 + s, \frac{\lambda}{K}\right) = \pi K\left(\frac{1}{2} - \frac{\mu}{K}\right) + O(s)$$

in analogy with (4.23). Similarly (4.21) shows that

$$\begin{aligned}\lim_{s\to 0} \sum_{\nu=1}^{k} \sin 2\pi \frac{\mu^*\nu}{k} \cdot \zeta\left(1 - s, \frac{\nu}{k}\right) &= \pi k\left(\frac{1}{2} - \frac{\mu^*}{k}\right) + \frac{\pi k}{2}\, \delta\left(\frac{\mu^*}{k}\right)\\ &= \pi k\left\{\frac{1}{2} - \frac{\mu^*}{k} + \left(\left[\frac{\mu^*}{k}\right] - \delta\left(\frac{\mu^*}{k}\right)\right) + \frac{1}{2}\delta\left(\frac{\mu^*}{k}\right)\right\}\\ &= \pi k\left\{\frac{1}{2} - \frac{\mu^*}{k} + \left[\frac{\mu^*}{k}\right] - \frac{1}{2}\delta\left(\frac{\mu^*}{k}\right)\right\} = -\pi k\left(\left(\frac{\mu^*}{k}\right)\right),\end{aligned}$$

the last step being obtained by the aid of (4.52). From (5.3) we infer that

$$\mu^* \equiv -h\mu_a \qquad (\text{mod } k);$$

thus

$$\left(\left(\frac{\mu^*}{k}\right)\right) = -\left(\left(-\frac{h\mu}{k}\right)\right).$$

The last few results give

$$R_2 = -\tfrac{1}{2}\sum_{\mu_a}\left(\left(\frac{\mu}{K}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right) = -\tfrac{1}{2}t_a(h, k). \tag{6.5}$$

From (5.42), (5.7), (6.1), (6.2), (6.3), and (6.5) we obtain the transformation equation of $F_a(x)$ for $5 \nmid k$. We combine this result with (4.7) in the following

THEOREM 1. *The function $F_a(x)$, defined by* (1.1), *satisfies the transformation equation*

$$\begin{aligned} F_a\left(\exp\left[2\pi i\frac{h}{k} - 2\pi\frac{z}{k}\right]\right) &= \omega_a(h, k)\exp\left[\frac{\pi}{30k}\left(\frac{B}{z} - Az\right)\right] \\ &\quad \cdot F_b\left(\exp\left[2\pi i\frac{h'}{k} - \frac{2\pi}{kz}\right]\right) \end{aligned} \tag{6.6}$$

for $k \equiv 0 \pmod 5$ *and the equation*

$$\begin{aligned} &F_a\left(\exp\left[2\pi i\frac{h}{k} - 2\pi\frac{z}{k}\right]\right) \\ &= \frac{1}{2}\chi_a(h, k)\csc\frac{\pi\alpha}{5}\exp\left[\frac{\pi}{30k}\left(\frac{1}{z} - Az\right)\right]H_a\left(\exp\left[2\pi i\frac{H'}{k} - \frac{2\pi}{Kz}\right]\right) \end{aligned} \tag{6.7}$$

for $k \not\equiv 0 \pmod 5$.

References to the meaning of the symbols in (6.6) are given in (4.7). In (6.7), H_a is defined by (5.8), H' and K by (1.3), and $\chi_a(h, k)$ by

$$\chi_a(h, k) = \exp\left(\pi i t_a(h, k)\right), \tag{6.71}$$

with the $t_a(h, k)$ of (6.5).

7. We can translate the foregoing results into the language of modular functions. Let τ be a complex variable where $\Im\tau > 0$. Setting

$$x = \exp(2\pi i\tau), \qquad x' = \exp(2\pi i\tau'), \tag{7.11}$$

we have by comparison with (1.2)

$$\tau = \frac{h}{k} + \frac{iz}{k}, \qquad \tau' = \frac{h'}{k} + \frac{i}{kz}, \tag{7.12}$$

so that

$$z = \frac{1}{i}(k\tau - h), \qquad \frac{1}{z} = \frac{1}{i}(k\tau' - h'). \tag{7.13}$$

This is consistent since $\Re z > 0$ (see (1.2)) insures $\Im\tau > 0$ in (7.12).

Let $V\tau$ be a modular substitution:

$$V\tau = \tau' = \frac{\alpha\tau + \beta}{\gamma\tau + \delta}, \qquad \alpha\delta - \beta\gamma = 1; \alpha, \beta, \gamma, \delta \text{ integers.}$$

Then clearly

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} h' & -(hh' + 1)k^{-1} \\ k & -h \end{pmatrix} \tag{7.2}$$

is a modular substitution which transforms τ into τ', as we see from (7.12).

Now consider the congruence subgroup $H(5)$ of the full modular group $\Gamma(1)$ given by the condition

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \equiv \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \pmod 5. \tag{7.3}$$

$H(5)$ is of "level" 5 ("Stufe" in the terminology of F. Klein) and is of index 12 in $\Gamma(1)$. First, if $\gamma = 0$, then evidently

$$\tau' = \tau + \beta,$$

and from (1.1)

$$F_a(\exp[2\pi i\tau]) = F_a(\exp[2\pi i\tau']).$$

If $\gamma \neq 0$ we take $\gamma > 0$, and have from (6.6), (7.13)

$$F_a(\exp(2\pi i\tau)) \exp\left(-\frac{\pi i A\tau}{30}\right)$$
$$= \omega_a(-\delta, \gamma) \exp\left\{\frac{\pi i A}{30\gamma}(\alpha + \delta)\right\} F_a(\exp(2\pi i\tau')) \exp\left(-\frac{\pi i A\tau'}{30}\right),$$

since $h \equiv 1 \pmod 5$ implies $b = a$, $B = A$.

We define

$$f_a(\tau) = F_a(\exp(2\pi i\tau)) \exp\left(-\frac{\pi i A\tau}{30}\right) = i\,\frac{\eta(5\tau)}{\vartheta_1(a\tau \mid 5\tau)} \exp\left(-\frac{\pi i a^2\tau}{5}\right); \tag{7.4}$$

the last equality subsists by virtue of (2.1), (3.43). We define also the *multiplier*

$$M(\alpha, \beta, \gamma, \delta) = \begin{cases} 1, & \gamma = 0, \\ \omega_a(\delta, \gamma) \exp\left\{-\frac{\pi i A}{30\gamma}(\alpha + \delta)\right\}, & \gamma > 0. \end{cases} \tag{7.5}$$

Then we have

THEOREM 1a. *The functions*

$$f_a(\tau) = i\,\frac{\eta(5\tau)}{\vartheta_1(a\tau\,|\,5\tau)}\exp\left(-\frac{\pi i a^2\tau}{5}\right) \qquad (a = 1, 2),$$

are modular functions associated with the subgroup $H(5)$ defined by (7.3) and the multiplier $M(\alpha, \beta, \gamma, \delta)$ of (7.5). That is, their transformation equation is

$$f_a(V\tau) = M(\alpha, \beta, \gamma, \delta)f_a(\tau), \qquad V\tau \in H(5). \tag{7.6}$$

It remains to discuss the behavior of $f_a(\tau)$ under those substitutions $V\tau$ which are in $\Gamma(1)$ but not in $H(5)$. First, let $\gamma \equiv 0 \pmod 5$. This condition defines the well-known congruence subgroup usually denoted by $\Gamma_0(5)$. We cannot have $\gamma = 0$ since this enforces $\alpha \equiv \delta \equiv 1 \pmod 5$. Hence, from (6.6), we get

$$f_a(\tau) = \omega_a(-\delta, \gamma)\exp\left\{\frac{\pi i}{30\gamma}(B\alpha + A\delta)\right\}f_b(\tau'). \tag{7.7}$$

If $\gamma \not\equiv 0 \pmod 5$ we need to use the transformation equation (6.7) instead of (6.6). The functions $H_\alpha(x)$ of (5.8) are readily expressed in terms of ϑ-functions:

$$H_\alpha(x) = 2\sin\frac{\pi\alpha}{5}\exp\left(\frac{\pi i\tau}{6}\right)\frac{\eta(\tau)}{\vartheta_1(\alpha/5\,|\,\tau)}. \tag{7.81}$$

Also

$$x'' = \exp\left\{2\pi i\left(\frac{H'}{k} + \frac{i}{Kz}\right)\right\} = \exp\left(2\pi i\frac{\tau'}{5}\right), \tag{7.82}$$

since $5H' \equiv h' \pmod k$. Using (7.81), (7.82), (2.1) in (6.7), we readily derive:

$$\begin{aligned}&\exp\left(-\frac{\pi i a^2\tau}{5}\right)\frac{\eta(5\tau)}{\vartheta_1(a\tau\,|\,5\tau)}\\ &\quad= \frac{1}{i}\chi_a(-\delta, \gamma)\exp\left(\frac{\pi i}{30\gamma}(\alpha + B\delta)\right)\frac{\eta(\tau'/5)}{\vartheta_1(\sigma/5\,|\,\tau'/5)}, \qquad \gamma\sigma \equiv a \pmod 5.\end{aligned} \tag{7.91}$$

II. The Kloosterman sums

8. In this part we discuss the order of magnitude of certain sums of the roots of unity $\omega_a(h, k)$, $\chi_a(h, k)$, which occur in the transformation equations (6.6), (6.7). We prove that these sums are in fact "incomplete Kloosterman sums" and are therefore subject to the estimate $O(n^{1/3}k^{2/3+\epsilon})$ rather than the trivial $O(k)$. The first step (§§8, 9) will be to reduce the quantities $\omega_a(h, k)$ and $\chi_a(h, k)$ to an explicit exponential form.

We consider $\omega_a(h, k)$ first and note from its definition (4.71) that we need to determine $r_a(h, k)$ modulo 2. It will turn out that it is convenient to discuss $30kr_a(h, k)$ modulo $60k$. First,

$$r_a(h, k) = \sum_{\mu_a}\left(\left(\frac{\mu}{k}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right) = \sum_{\mu_a}\left(\frac{\mu}{k} - \frac{1}{2}\right)\left(\left(\frac{h\mu}{k}\right)\right).$$

But

$$\sum_{\mu_a}\left(\left(\frac{h\mu}{k}\right)\right) = \sum_{\mu_a}\left(\left(-\frac{h\mu}{k}\right)\right) = -\sum_{\mu_a}\left(\left(\frac{h\mu}{k}\right)\right) = 0,$$

by virtue of (4.54), (4.53). Hence,

$$r_a(h, k) = \sum_{\mu_a}\frac{\mu}{k}\left(\left(\frac{h\mu}{k}\right)\right). \tag{8.1}$$

Next, we define $\mathbf{M}_a$ to be the system of values

$$\mathbf{M}_a \equiv a \pmod 5, \qquad 0 < \mathbf{M}_a < k. \tag{8.2}$$

Then

$$r_a(h, k) = \sum_{\mathbf{M}_a}\frac{\mu}{k}\left(\left(\frac{h\mu}{k}\right)\right) + \sum \frac{k-\mu}{k}\cdot\left(\left(\frac{h(k-\mu)}{k}\right)\right)$$

$$= 2\sum_{\mathbf{M}_a}\frac{\mu}{k}\left(\left(\frac{h\mu}{k}\right)\right) + \sum_{\mathbf{M}_a}\left(\left(-\frac{h\mu}{k}\right)\right).$$

Now

$$\sum_{\mathbf{M}_a}\left(\left(-\frac{h\mu}{k}\right)\right) = -\sum_{\mathbf{M}_a}\left(\left(\frac{h\mu}{k}\right)\right) = -\sum_{\lambda=0}^{k_1-1}\left(\left(\frac{5h\lambda + ha}{k}\right)\right)$$

$$= -\sum_{\lambda \bmod k_1}\left(\left(\frac{h\lambda}{k_1} + \frac{ha}{k}\right)\right) = -\left(\left(\frac{ha}{5}\right)\right),^{7} \qquad k_1 = \frac{k}{5};$$

also

$$\left(\left(\frac{ha}{5}\right)\right) = \frac{ha}{5} - \left[\frac{ha}{5}\right] - \frac{1}{2} = \frac{c}{5} - \frac{1}{2},$$

where

$$c = \{ha, 5\}. \tag{8.21}$$

Therefore

$$r_a(h, k) = 2\sum_{\mathbf{M}_a}\frac{\mu}{k}\left(\left(\frac{h\mu}{k}\right)\right) - \left(\frac{c}{5} - \frac{1}{2}\right). \tag{8.22}$$

Using the definition (4.51) of $((x))$, we get from (8.22),

$$30kr_a(h, k) = 60\frac{h}{k}\sum_{\mathbf{M}_a}\mu^2 - 60\sum_{\mathbf{M}_a}\mu\left[\frac{h\mu}{k}\right] - 30\sum_{\mathbf{M}_a}\mu - 3k(2c-5)$$

$$= 2h(2k^2 + 3k(2a-5) + A) \tag{8.23}$$

$$- 3k(k + 2a + 2c - 10) - 60\sum_{\mathbf{M}_a}\mu\left[\frac{h\mu}{k}\right],$$

[7] For the last identity see [5], formula (2.4).

with the A of (3.43). (8.23) shows that $30kr_a(h, k)$ is always an integer. Furthermore

$$30kr_a(h, k) \equiv 0 \pmod 3, \text{ if } 3 \nmid k; \tag{8.31}$$

$$30kr_a(h, k) \equiv 2a + 2c - 1 \pmod 4, \qquad \text{if } k \text{ is odd.} \tag{8.32}$$

We shall now deduce a formula for $r_a(h, k)$ which will enable us to determine very easily its residue modulo various multiples of k. To do this we calculate a certain sum in two ways.[8] First,

$$\begin{aligned}\sum_{M_a}\left(\left(\frac{h\mu}{k}\right)\right)^2 &= \sum_{M_a}\left(\left(\frac{\mu}{k}\right)\right)^2 = \sum_{M_a}\left(\frac{\mu}{k} - \frac{1}{2}\right)^2 \\ &= \frac{1}{30k}(-k^2 + C) + \frac{1}{4}\sum_{M_a} 1, \qquad C = 6c^2 - 30c + 25,\end{aligned} \tag{8.41}$$

with c defined as in (8.21). On the other hand,

$$\begin{aligned}\sum_{M_a}\left(\left(\frac{h\mu}{k}\right)\right)^2 &= \sum_{M_a}\left(\frac{h\mu}{k} - \left[\frac{h\mu}{k}\right] - \frac{1}{2}\right)^2 \\ &= 2h\sum_{M_a}\frac{\mu}{k}\left(\frac{h\mu}{k} - \left[\frac{h\mu}{k}\right] - \frac{1}{2}\right) - \frac{h^2}{k^2}\sum_{M_a}\mu^2 \\ &\qquad + \sum_{M_a}\left[\frac{h\mu}{k}\right]\left(\left[\frac{h\mu}{k}\right] + 1\right) + \frac{1}{4}\sum_{M_a} 1 \\ &= hr_a(h, k) + h\left(\frac{c}{5} - \frac{1}{2}\right) \\ &\qquad - \frac{h^2}{30k}(2k^2 + 3k(2a - 5) + A) + 2S + \frac{1}{4}\sum_{M_a} 1,\end{aligned} \tag{8.42}$$

where we have made use of (8.22). S is an integer defined by

$$S = \frac{1}{2}\sum_{M_a}\left[\frac{h\mu}{k}\right]\left(\left[\frac{h\mu}{k}\right] + 1\right).$$

Comparison of (8.41) and (8.42) yields

$$\begin{aligned}&30kr_a(h, k) \\ &\quad = h^2(2k^2 + 3k(2a - 5) + A) - (k^2 - C) - 3hk(2c - 5) - 60kS.\end{aligned} \tag{8.5}$$

Let $60 = fG$, where f is the greatest divisor of 60 which is prime to k. There are four cases:

$$\begin{aligned}&(k, 60) = 5, f = 12, G = 5; && (k, 60) = 10, 20, f = 3, G = 20; \\ &(k, 60) = 15, f = 4, G = 15; && (k, 60) = 30, 60, f = 1, G = 60.\end{aligned} \tag{8.61}$$

[8] Cf. [5], proof of Theorem 3.

We write h' to denote a solution of

(8.62) $$hh' \equiv -1 \pmod{Gk}.$$

This congruence is always solvable since all primes in G divide k and therefore $(h, k) = 1$ implies $(h, Gk) = 1$.

From (8.5) we obtain by multiplication with $-h'$

(8.71) $$\begin{gathered} 30kr_a(h, k) \equiv uh - vh' - 3k(2c - 5) \pmod{Gk}, \\ u = 3k(2a - 5) + A, \qquad v = -k^2 + C, \end{gathered}$$

since $Gk \mid 2k^2$. Moreover, examination of (8.61), (8.31), (8.32) yields

(8.72) $$30kr_a(h, k) \equiv 6(a + c) + 3 \pmod{f}.$$

Now, setting

(8.73) $$\begin{aligned} f\varphi &\equiv 1 \pmod{Gk}, \\ Gk\Gamma &\equiv 1 \pmod{f}, \end{aligned}$$

we have from (8.71), (8.72)

(8.8) $$\begin{aligned} &30kr_a(h, k) \\ &\quad\equiv f\varphi(uh - vh' - 3k(2c - 5)) + Gk\Gamma(6(a + c) + 3) \pmod{60k}. \end{aligned}$$

Hence, finally,

(8.9) $$\begin{aligned} \omega_a(h, k) &= \exp\left(\frac{2\pi i}{60k}\cdot 30kr_a(h, k)\right) \\ &= \exp\left\{2\pi i\left(\frac{\Gamma}{f}(6a + 6c + 3) - \frac{3\varphi}{G}(2c - 5) + \frac{\varphi}{Gk}(uh - vh')\right)\right\}. \end{aligned}$$

9. We now turn our attention to the root of unity $\chi_a(h, k)$. Since the treatment here is quite similar to that of the preceding section, we shall merely display the formulas without much explanation. Corresponding formulas will be denoted by the same decimal part, i.e., (9.42) is the analogue of (8.42).[9]

(9.1) $$\begin{aligned} t_a(h, k) &= \sum_{\mu_a}\left(\left(\frac{\mu}{K}\right)\right)\left(\left(\frac{h\mu}{k}\right)\right) \\ &= \sum_{\mu_a}\left(\frac{\mu}{K} - \frac{1}{2}\right)\left(\left(\frac{h\mu}{k}\right)\right) = \sum \frac{\mu}{K}\left(\left(\frac{h\mu}{k}\right)\right). \end{aligned}$$

(9.2) $$\mathrm{M}_a \equiv a \pmod 5, \qquad 0 < \mathrm{M}_a < K.$$

[9] See footnote 6. We shall write μ_a for $\mu_a(K)$ throughout this section. Also, $\mathrm{M}_a = \mathrm{M}_a(K)$ will have the meaning of (8.2) but with the modulus K (see (9.2)). In this section frequent use will be made of the remark in the lines preceding (5.6).

$$t_a(h, k) = \sum_{M_a} \frac{\mu}{K}\left(\left(\frac{h\mu}{k}\right)\right) + \sum_{M_a} \frac{K-\mu}{K}\cdot\left(\left(\frac{h(K-\mu)}{k}\right)\right)$$

(9.22)
$$= 2\sum_{M_a} \frac{\mu}{K}\left(\left(\frac{h\mu}{k}\right)\right) + \sum_{M_a}\left(\left(-\frac{h\mu}{k}\right)\right) = 2\sum_{M_a}\frac{\mu}{K}\left(\left(\frac{h\mu}{k}\right)\right).$$

$$t_a(h, k) = 2\sum \frac{\mu}{K}\left(\frac{h\mu}{k} - \left[\frac{h\mu}{k}\right] - \frac{1}{2} + \frac{1}{2}\,\delta\left(\frac{h\mu}{k}\right)\right),$$

(9.23)
$$30kt_a(h, k) = 2h(2K^2 + 3K(2a-5) + A) - 3k(K + 2a - 2\alpha - 5) - 12\sum_{M_a}\mu\left[\frac{h\mu}{k}\right],$$

with α as defined in (5.6).

(9.31) $$30kt_a(h, k) \equiv 0 \pmod 3, \text{ if } 3 \nmid k;$$

(9.32) $$30kt_a(h, k) \equiv 3k + 2a + 2\alpha + 1 \pmod 4 \text{ if } k \text{ is odd.}$$

The discussion for the modulus 5 has no analogue in §8. From (9.2) we derive

(9.33)
$$30kt_a(h, k) \equiv 2Ah - k(a - \alpha) - 2\sum_{M_a} a\left[\frac{h\mu}{k}\right] \pmod 5$$
$$\equiv 2a^2h - a(k-1) - 2a\sum_{M_a}\left[\frac{h\mu}{k}\right] \pmod 5.$$

Now

$$\sum_{M_a}\left[\frac{h\mu}{k}\right] = \sum_{\lambda=0}^{k-1}\left[\frac{5\lambda h + ah}{k}\right]$$
$$= -\sum_{\lambda \bmod k}\left(\left(\frac{5\lambda h + ah}{k}\right)\right) + \sum_{\lambda=0}^{k-1}\frac{5\lambda h + ah}{k} - \sum_{0}^{k-1}\frac{1}{2} + \frac{1}{2}$$
$$= \frac{h}{2}(5k + 2a - 5) - \frac{k-1}{2}.$$

Hence,

$$2a\sum_{M_a}\left[\frac{h\mu}{k}\right] \equiv 2a^2h - a(k-1) \pmod 5$$

and from (9.33) we obtain

(9.34) $$30kt_a(h, k) \equiv 0 \pmod 5.$$

(9.5)
$$30hkt_a(h, k) = h^2(2K^2 + 3K(2a-5) + A) - (k^2 - 1) + 6hk\alpha - 12kS, \; S \text{ integral.}$$

Let $60 = Fg$, where F is the greatest divisor of 60 prime to k.

$$\begin{aligned}(9.61)\qquad &(k, 60) = 1, F = 60, g = 1; \qquad (k, 60) = 2,4, \;\; F = 15, g = \;\;4;\\ &(k, 60) = 3, F = 20, g = 3; \qquad (k, 60) = 6,12, F = \;\;5, g = 12.\end{aligned}$$

$$(9.62)\qquad hh' \equiv -1 \pmod{gk}.$$

$$\begin{aligned}(9.71)\qquad 30kt_a(h, k) &\equiv uh - vh' + 6k\alpha \pmod{gk},\\ u = 3K(2a - 5) + A, &\qquad v = -k^2 + 1.\end{aligned}$$

$$(9.72)\qquad 30kt_a(h, k) \equiv 15(k + 2a + 2\alpha + 3) \pmod{F}.$$

$$\begin{aligned}(9.73)\qquad F\Phi &\equiv 1 \pmod{gk},\\ gk\gamma &\equiv 1 \pmod{F}.\end{aligned}$$

$$\begin{aligned}(9.8)\qquad &30kt_a(h, k)\\ &\equiv F\Phi(uh - vh' + 6k\alpha) + gk\gamma\cdot 15(k + 2a + 2\alpha + 3) \pmod{60k}.\end{aligned}$$

$$\chi_a(h, k) = \exp\left\{2\pi i\left(\frac{6\Phi\alpha}{g} + \frac{15\gamma}{F}(k + 2a + 2\alpha + 3) + \frac{\Phi}{gk}(uh - vh')\right)\right\}.$$

Since $5H' \equiv h' \pmod{gk}$, $G = 5g$, the last formula becomes

$$\begin{aligned}(9.9)\qquad &\chi_a(h, k)\\ &= \exp\left\{2\pi i\left(\frac{6\Phi\alpha}{g} + \frac{15\gamma}{F}(k + 2a + 2\alpha + 3) + \frac{\Phi}{Gk}(uH - 25vH')\right)\right\}.\end{aligned}$$

10. We are now ready to consider the following sum:[10]

$$\begin{aligned}(10.1)\qquad &A(n, \nu; k; d; \sigma_1, \sigma_2; a) = T = \sum_h{}' \omega_a(h, k) \exp\left(-\frac{2\pi i}{k}(hn - h'\nu)\right);\\ &h \bmod k, h \equiv d \pmod 5; \sigma_1 \leqq h' < \sigma_2; 0 \leqq \sigma_1 < \sigma_2 \leqq k, 5 \mid k, 5 \nmid d.\end{aligned}$$

The notation $\sum'$ means that h runs over integers prime to M, where M is the modulus of the sum. h' denotes any solution of

$$(10.2)\qquad hh' \equiv -1 \pmod{M}.$$

These meanings will persist to the end of the paper. The restriction on h' is to be interpreted modulo k.

The expression (8.9) for $\omega_a(h, k)$ does not exhibit its true periodicity, but from (4.71) we see that it has the period k. Thus we have, using (8.9),

[10] Hereafter, the summation conditions will be placed at the end of the display line or in the text, rather than underneath the summation sign.

$$T = \epsilon(a, k, d) \cdot \frac{1}{G} \sum_s{}' \exp\left\{\frac{2\pi i}{Gk} f(s)\right\}, \tag{10.31}$$
$$s \bmod Gk;\ s \equiv d \pmod 5,\ \sigma_1 \leqq s' < \sigma_2\ ;\ |\ \epsilon(a, k, d)\ | = 1,$$

where

$$f(s, s'; u, v, n) = f(s) = (\varphi u - Gn)s - (\varphi v - G\nu)s'. \tag{10.32}$$

In (10.31) s' is restricted to G (or possibly $G + 1$) intervals whose endpoints are congruent modulo k to σ_1, σ_2. More accurately, the condition on s' would be expressed as follows: $0 \leqq s' < Gk$, $\sigma_1 \leqq \{s', k\} < \sigma_2$.

Let[11] $m(s)$ be defined in $(0, k)$ by

$$m(s) = \begin{cases} 1, & \text{for } \sigma_1 \leqq s < \sigma_2, \\ 0, & \text{elsewhere in the interval } 0 \leqq s < k, \end{cases} \tag{10.41}$$

and outside of this interval by periodicity. Then

$$m(s) = \sum_{l=0}^{k-1} \alpha_l \exp\left(2\pi i \frac{sl}{k}\right),$$

with

$$\alpha_j = \frac{1}{k} \sum_{s=0}^{k-1} m(s) \exp\left(-2\pi i \frac{sj}{k}\right) = \frac{1}{k} \sum_{s=\sigma_1}^{\sigma_2-1} \exp\left(-\frac{2\pi isj}{k}\right),$$

$$\alpha_0 = \frac{\sigma_2 - \sigma_1}{k}, \qquad |\ \alpha_0\ | \leqq 1,$$

while for $j \neq 0$,

$$|\ \alpha_j\ | \leqq \frac{2}{k} \left| 1 - \exp\left(-\frac{2\pi ij}{k}\right)\right|^{-1} = \frac{1}{k} \csc \frac{\pi j}{k}.$$

Hence,

$$\sum_{j=0}^{k-1} |\ \alpha_j\ | \leqq 1 + \sum_{j=1}^{\frac{1}{2}k} \frac{1}{j} = O(\log k). \tag{10.42}$$

This gives

$$\begin{aligned} T &= O\left(\sum_s{}' m(s') \exp\left\{\frac{2\pi i}{Gk} f(s)\right\}\right) \\ &= O\left(\sum_{l=0}^{k-1} \alpha_l \sum_s{}' \exp\left\{\frac{2\pi i}{Gk} (f(s) + s'l)\right\}\right), \qquad s \bmod Gk,\ s \equiv d \pmod 5. \end{aligned} \tag{10.43}$$

The condition $s \equiv d \pmod 5$ on the inner sum of (10.43) can be easily removed. We note that

$$\frac{1}{5} \sum_{r=1}^{5} \exp\left(2\pi i \frac{r(s' - d')}{5}\right) = \begin{cases} 1, & s' \equiv d' \pmod 5, \\ 0, & \text{otherwise}. \end{cases} \tag{10.51}$$

[11] This device is taken from [1], p. 94.

Since $(d, 5) = 1$ the condition $s \equiv d \pmod 5$ is equivalent to $s' \equiv d' \pmod 5$. Using (10.51) we find that the inner sum of (10.43) is equal to

$$\sum_s{}' \exp\left\{\frac{2\pi i}{Gk}\,(f(s) + s'l)\right\} \frac{1}{5}\sum_{r=1}^{5} \exp\left\{2\pi i\,\frac{r(s' - d')}{5}\right\}$$

$$= \frac{1}{5}\sum_{r=1}^{5} \exp\left(-\,2\pi i\,\frac{rd'}{5}\right) \sum_s{}' \exp\left\{\frac{2\pi i}{Gk}\,(f(s) + s'(l + rgk))\right\} \tag{10.52}$$

$$= O\left(\sum_s{}' \exp\left\{\frac{2\pi i}{Gk}\,((\varphi u - Gn)s - (\varphi v - G\nu - l - rgk)s'\right\}\right), \qquad s \bmod Gk.$$

The sum in the extreme right member of (10.52) is a complete Kloosterman sum. We can therefore make use of an estimate of Salié ([6], p. 264), so that from (10.42), (10.43), (10.52)

$$T = O((\varphi u - Gn, Gk)^{\frac{1}{3}}\cdot(Gk)^{\frac{2}{3}+\epsilon}\log k).$$

But from the meaning (8.71), (8.73) of u, f, and G,

$$(\varphi u - Gn, Gk) = (f\varphi u - fGn, Gk) = (u - 60n, Gk)$$
$$= (A - 60n, k) = O(n), \qquad n \geqq 1;$$

hence

$$T = O(n^{\frac{1}{3}}k^{\frac{2}{3}+\epsilon}). \tag{10.6}$$

The preceding discussion justifies the following

THEOREM 2. *The sum*

$$A(n, \nu; k; d; \sigma_1, \sigma_2; a) = \sum_h{}' \omega_a(h, k) \exp\left(-\,2\pi i\,\frac{hn - h'\nu}{k}\right); \tag{10.71}$$
$$h \bmod k,\ h \equiv d(5),\ \sigma_1 \leqq h' < \sigma_2$$

in which the parameters are all integers, $n > 0, k > 0, 5 \mid k, 5 \nmid d, 0 \leqq \sigma_1 < \sigma_2 \leqq k$. and $a = 1$ or 2, is subject to the estimate

$$O(n^{\frac{1}{3}}k^{\frac{2}{3}+\epsilon}) \tag{10.72}$$

uniformly in $\nu, d, \sigma_1, \sigma_2, a$.

Quite similar considerations apply to sums involving $\chi_a(h, k)$. We readily derive

THEOREM 3. *The sum*

$$B(n, \nu; k; \sigma_1, \sigma_2; a) = \sum_h{}' \chi_a(h, k) \exp\left(-\,2\pi i\,\frac{hn - H'\nu}{k}\right); \tag{10.81}$$
$$h \bmod k,\ HH' \equiv -\,1 \pmod k,\ \sigma_1 \leqq h' < \sigma_2,$$

has the estimate (10.72) *uniformly in $\nu, \sigma_1, \sigma_2, a$.*

We can carry through the proof of Theorem 3 by using H and H' rather than h and h'. This is permissible since $HH' \equiv -1 \pmod{k}$. The sum is again transformed into one with the modulus Gk by making use of (9.9) and

$$\frac{hn - H'\nu}{k} = \frac{Ghn - GH'\nu}{Gk} = \frac{Hgn - H'G\nu}{Gk}.$$

The result then follows as before.

III. A convergent series for $p_a(n)$: Asymptotic formulas

11. The final step is to apply the Hardy-Littlewood method, in the sharper form developed by Kloosterman and Rademacher, to the generating functions $F_a(x)$ of (1.1). Since this method has been exposed in complete detail at least three times,[12] it does not seem necessary or desirable to present it again here, especially in view of the extreme typographical complexity of the formulas. We give only a synopsis.

The notations used are those of Rademacher [3] and references to that paper will be enclosed in braces. First, we have {(3.1), (4.2)}

$$\begin{aligned} p_a(n) &= \frac{1}{2\pi i}\int_C \frac{F_a(x)}{x^{n+1}}\,dx \\ &= \sum_{h,k}{}' \exp\left(-2\pi i\frac{hn}{k}\right)\int_{-\vartheta'}^{\vartheta''} F_a\left(\exp\left[2\pi i\frac{h}{k} - 2\pi\frac{z}{k}\right]\right)\exp 2\pi nw\,d\varphi; \qquad (11.1)\\ &\qquad n \geqq 1;\ 0 \leqq h < k \leqq N,\ \vartheta' = \vartheta'_{h,k},\ \vartheta'' = \vartheta''_{h,k}. \end{aligned}$$

To (11.1) we apply the transformation formulas of Theorem 1, but we must distinguish several cases according to the residues of h and k modulo 5. Accordingly we put

$$p_a(n) = p_a^{(1)}(n) + p_a^{(2)} = \sum_{d=1}^{4} p_{a,d}(n) + p_a^{(2)}(n), \tag{11.2}$$

where $p_{a,d}^{(1)}(n)$ is the sum of those terms in the extreme right member of (11.1) for which $5 \mid k$, $h \equiv d \pmod 5$, while $p_a^{(2)}(n)$ is the same sum for all k with $5 \nmid k$.

Thus we obtain

$$\begin{aligned} p_{a,d}^{(1)}(n) &= \sum_{h,k}{}' \omega_a(h,k)\exp\left(-2\pi i\frac{hn}{k}\right)\int_{-\vartheta'}^{\vartheta''}\exp\left(\frac{\pi B}{30\,k^2 w} + \pi w\left(2n - \frac{A}{30}\right)\right)d\varphi \\ &\quad + \sum_{h,k}{}' \omega_a(h,k)\exp\left(-2\pi i\frac{hn}{k}\right)\int_{-\vartheta'}^{\vartheta''}\sum_{\nu=1}^{\infty} p_b(\nu)\exp\left(2\pi i\frac{h'\nu}{k}\right) \qquad (11.3)\\ &\quad \cdot\exp\left(-\frac{\pi}{k^2 w}\left(2\nu - \frac{B}{30}\right) + \pi w\left(2n - \frac{A}{30}\right)\right)d\varphi; \\ &\qquad 5 \mid k,\ 0 \leqq h < k \leqq N,\ h \equiv d \pmod 5. \end{aligned}$$

[12] Namely, Rademacher [3]; I. Niven, *On a certain partition function*, American Journal of Mathematics, vol. 62(1940), pp. 353–364; M. Haberzetle, *On some partition functions*, American Journal of Mathematics, vol. 63(1941), pp. 589–599.

Since the coefficient of w^{-1} in the exponent of the integrand of the second sum is always negative, this sum will furnish no contribution to the final result, as we see from §7 of Rademacher's paper. The same will be true of the first sum if $B = -11$ (see (3.6)), but if $B = 1$, the first sum will be responsible for a principal term {§5}

$$\frac{2\pi}{(60n - A)^{\frac{1}{2}}} \sum_k \frac{A_{k,d}(n)}{k} I_1\left(\frac{\pi(60n - A)^{\frac{1}{2}}}{15k}\right), \qquad 1 < k \leqq N, 5 \mid k, \tag{11.41}$$

and an error term {§§6–7}

$$O(n^{\frac{1}{2}} \exp(3\pi n N^{-2}) \cdot N^{-\frac{1}{2}+\epsilon}), \tag{11.42}$$

where

$$A_{k,d}(n) = A(n, 0; k; d; 0, k; a).$$

Note that the estimates for the Kloosterman sums{(5.3), (6.2)} have been prepared in our Theorem 2.[13] Moreover, $B = 1$ implies $b = 1$, which in turn enforces $h \equiv \pm a \pmod 5$, i.e., $d \equiv \pm a \pmod 5$ by (2.42). Thus, using (11.2) we see that the principal term of $p_a^{(1)}(n)$ will be

$$\frac{2\pi}{(60n - A)^{\frac{1}{2}}} \sum_k \frac{A_k(n)}{k} I_1\left(\frac{\pi(60n - A)^{\frac{1}{2}}}{15k}\right), \qquad 1 \leqq k \leqq N, 5 \mid k, \tag{11.5}$$

with

$$A_k(n) = \sum_h{}' \omega_a(h, k) \exp\left(-2\pi i \frac{hn}{k}\right), \quad 0 \leqq h < k, h \equiv \pm a \pmod 5). \tag{11.51}$$

If we go back to (11.1) we find

$$\begin{aligned} p_a^{(2)}(n) = {} & \frac{1}{2} \sum_{h,k}{}' \csc \frac{\pi\alpha}{5} \chi_a(h, k) \\ & \cdot \exp\left(-2\pi i \frac{hn}{k}\right) \int_{\vartheta'}^{\vartheta''} \exp\left(\frac{\pi}{30k^2 w} + \pi w\left(2n - \frac{A}{30}\right)\right) d\varphi \\ & + \frac{1}{2} \sum_{h,k}{}' \csc \frac{\pi\alpha}{5} \chi_a(h, k) \exp\left(-2\pi i \frac{hn}{k}\right) \int_{\vartheta'}^{\vartheta''} \sum_{\nu=1}^{\infty} c_a(\nu) \exp\left(2\pi i \frac{H'\nu}{k}\right) \\ & \cdot \exp\left(-\frac{\pi}{Kkw}\left(2\nu - \frac{1}{6}\right) + \pi w\left(2n - \frac{A}{30}\right)\right) d\varphi, \quad 0 \leqq h < k \leqq N, 5 \nmid k. \end{aligned} \tag{11.6}$$

We can follow the same method as before. The estimates for the Kloosterman sums have been prepared in Theorem 3.[13] Thus we obtain as the principal term for $p_a^{(2)}(n)$

$$\frac{\pi}{(60n - A)^{\frac{1}{2}}} \sum_k \csc \frac{\pi\alpha}{5} \cdot \frac{B_k(n)}{k} I_1\left(\frac{\pi(60n - A)^{\frac{1}{2}}}{15k}\right), \qquad 1 \leqq k \leqq N, 5 \nmid k, \tag{11.71}$$

[13] See the lines following (6.1) in [3]. In our case the analogous sum can be expressed as $A(n, 0; k; d; \sigma_1, \sigma_2; a)$ or as the sum of two such A's, where σ_1, σ_2 are suitably chosen. For the case $5 \nmid k$ we use the B sums of Theorem 3.

with the error term (11.42). Here

$$B_k(n) = \sum_h{}' \chi_a(h, k) \exp\left(-2\pi i \frac{hn}{k}\right), \qquad 0 \leqq h < k, 5 \nmid k. \tag{11.72}$$

In this result we remark that csc $(\pi\alpha/5)$ can be replaced by $|\csc(\pi ak/5)|$ since $k\alpha \equiv a \pmod 5$ and $k^2 \equiv \pm 1 \pmod 5$ imply $\alpha \equiv \pm ak \pmod 5$.

Now combining (11.2), (11.5), (11.71), (11.42), and letting $N \to \infty$ while n remains fixed, we obtain our main

THEOREM 4. *The number, $p_a(n)$, of partitions of a positive integer n into positive summands of the form $5l \pm a$ $(a = 1, 2)$ is given by the convergent series*

$$\begin{aligned} p_a(n) = {} & \frac{2\pi}{(60n - A)^{\frac12}} \sum_{\substack{k>0 \\ 5 \mid k}} \frac{A_k(n)}{k} I_1\left(\frac{\pi(60n - A)^{\frac12}}{15k}\right) \\ & + \frac{\pi}{(60n - A)^{\frac12}} \sum_{\substack{k>0 \\ 5 \nmid k}} \left|\csc \frac{\pi ak}{5}\right| \frac{B_k(n)}{k} I_1\left(\frac{\pi(60n - A)^{\frac12}}{15k}\right), \end{aligned} \tag{11.8}$$

where $A_k(n)$ is defined by (11.51), *$B_k(n)$ by* (11.72), *and $A = 1$ for $a = 1$, $A = -11$ for $a = 2$.*

12.[14] We can easily derive asymptotic formulas for $p_a(n)$ from (11.8). Indeed, with

$$t_a = t = \frac{\pi}{15}(60n - A)^{\frac12} \tag{12.1}$$

we have

$$\begin{aligned} p_a(n) &= \frac{\pi}{(60n - A)^{\frac12}} \csc \frac{\pi a}{5} I_1(t) \left\{1 + \sum_k \frac{B_k(n)}{k} \frac{|\csc(\pi k/5)|}{\csc(\pi a/5)} \frac{I_1(t/k)}{I_1(t)} \right. \\ &\qquad\qquad \left. + \sum_k \sin \frac{\pi a}{5} \frac{A_k(n)}{k} \frac{I_1(t/k)}{I_1(t)}\right\} \\ &= \frac{\pi}{(60n - A)^{\frac12}} \csc \frac{\pi a}{5} I_1(t) \{1 + S_1 + S_2\}, \end{aligned} \tag{12.2}$$

where in S_1, $5 \nmid k$, $k > 0$, and in S_2, $5 \mid k$, $k > 0$. We divide the summands in S_1 into two classes: $1 \leqq k \leqq [t]$, $k > [t]$; similarly for S_2. Then applying the asymptotic formulas[15]

$$I_1(z) = O(z), \qquad |z| < 1, \tag{12.31}$$

$$I_1(z) \sim e^z(2\pi z)^{-\frac12}, \qquad |z| > 1, \tag{12.32}$$

we find without difficulty

$$S_1, S_2 = O(n \exp[-cn^{\frac12}]), \qquad c = \text{const.},$$

[14] The idea of this section was suggested by Professor Rademacher.

[15] See [9], p. 203.

so that from (12.2) and a further application of (12.32) we derive

$$p_a(n) \sim \left(\frac{15}{2}\right)^{1/2} \frac{\csc(\pi a/5)}{(60n - A)^{3/4}} \exp\left\{\frac{\pi(60n - A)^{1/2}}{15}\right\} \qquad (a = 1, 2), \tag{12.4}$$

which is the desired asymptotic formula.

We now consider the ratio of $p_1(n)$ to $p_2(n)$. From (12.4) we get at once

$$\frac{p_1(n)}{p_2(n)} \sim \frac{\sin(2\pi/5)}{\sin(\pi/5)} = \left(\frac{5 + 5^{1/2}}{5 - 5^{1/2}}\right)^{1/2} = \frac{1 + 5^{1/2}}{2}. \tag{12.5}$$

The theorems of I. Schur referred to in §1 show that this difference in the magnitude of $p_1(n)$ and $p_2(n)$ is caused by those partitions of n into summands differing by at least two and containing the summand one. Moreover, the limit $\frac{1}{2}(1 + 5^{1/2})$ in (12.5) is in accordance with Schur's result ([7], p. 321, last paragraph) that

$$\frac{F_1(x)}{F_2(x)} \to \frac{1 + 5^{1/2}}{2} \qquad \text{as } x \to 1.$$

Calculation of $p_1(n)$ and $p_2(n)$ for $n \leqq 25$ indicates that the approach in (12.5) is usually from below. As a matter of fact $p_1(n)/p_2(n)$ is always less than $\frac{1}{2}(1 + 5^{1/2})$ from a certain point on, as is readily established by using the first two terms of the asymptotic expansion of $I_1(z)$.[15]

Bibliography

1. T. Estermann, *Vereinfachter Beweis eines Satzes von Kloosterman*, Abhandlungen Hamburgischen Mathematischen Seminar, vol. 7(1929), pp. 82–98.
2. H. Rademacher, *On the partition function p(n)*, Proceedings of the London Mathematical Society, (2), vol. 43(1937), pp. 241–254.
3. H. Rademacher, *The Fourier coefficients of the modular invariant J(τ)*, American Journal of Mathematics, vol. 60(1938), pp. 501–512.
4. H. Rademacher, *Zur Theorie der Modulfunktionen*, Journal für die reine und angewandte Mathematik, vol. 167(1931), pp. 312–336.
5. H. Rademacher and A. L. Whiteman, *Theorems on Dedekind sums*, American Journal of Mathematics, vol. 63(1941), pp. 377–407.
6. H. Salié, *Zur Abschätzung der Fourierkoeffizienten gänzer Modulformen*, Mathematische Zeitschrift, vol. 36(1933), pp. 263–278.
7. I. Schur, *Zur Additiven Zahlentheorie*, Sitzungsberichte der Berliner Akademie, 1926, pp. 488–495.
8. J. Tannery and J. Molk, *Élements de la Théorie des Fonctions Elliptiques*, Paris, 1896, vol. 2.
9. G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge, 1922.

University of Pennsylvania.

INEQUALITIES FOR TRIGONOMETRIC INTEGRALS

By Paul Civin

Introduction. We are concerned with the following problem: *If*

$$f(x) = \int_{-R}^{R} e^{ixt}\, ds(t), \tag{0.1}$$

$$g(x) = \int_{-R}^{R} \mu(t) e^{ixt}\, ds(t), \tag{0.2}$$

where $s(t)$ is a complex valued function of bounded variation on $[-R, R]$ and $\mu(t)$ is continuous, and if

$$|f(x)| \leqq M \qquad (-\infty < x < \infty), \tag{0.3}$$

then find a bound for $|g(x)|$.

The present paper is divided into two parts. In Part I, the general problem is considered for various classes of functions $\mu(t)$, sufficiently restricted so that non-trivial bounds for the corresponding functions can be found. Part II is devoted to the case in which $\mu(t) = (it)^{\alpha}$, $0 < \alpha < 1$, so that $g(x)$ is the fractional derivative of $f(x)$ of order α.

We shall consistently use the symbols appearing in (0.1), (0.2), and (0.3) with the meanings which they have in these formulas.

Part I

The following lemma gives the fundamental method of approach to the problem. We adopt the notation

$$p(n) = n\pi/R \qquad (n = 0, \pm 1, \pm 2, \cdots).$$

Lemma 1. *If*

$$\mu(t) e^{iat} = \sum_{-\infty}^{\infty} c_n e^{ip(n)t}, \tag{1.1}$$

where a is a real number, then

$$|g(x)| \leqq M \sum_{-\infty}^{\infty} |c_n| \qquad (-\infty < x < \infty). \tag{1.2}$$

Received May 21, 1941; presented to the American Mathematical Society (Preliminary report) April 26, 1940. The author wishes to thank Dr. R. P. Boas, Jr. for help in preparing this paper.

By the definition (0.2),

$$g(x) = \int_{-R}^{R} \mu(t) e^{ixt}\, ds(t)$$

$$= \int_{-R}^{R} \sum_{-\infty}^{\infty} \exp \{i[p(n) - a + x]t\}\, ds(t).$$

When $\sum_{-\infty}^{\infty} | c_n |$ is convergent, $\sum_{-\infty}^{\infty} c_n e^{ip(n)t}$ is uniformly convergent in $[-R, R]$, and the order of integration and summation may be changed. Relation (1.2) is trivial if $\sum_{-\infty}^{\infty} | c_n |$ does not converge. Hence, by the use of relation (0.1),

$$\begin{aligned} g(x) &= \sum_{-\infty}^{\infty} c_n \int_{-R}^{R} \exp \{i[p(n) - a + x]t\}\, ds(t) \\ &= \sum_{-\infty}^{\infty} c_n f[p(n) - a + x], \end{aligned} \tag{1.3}$$

and therefore

$$| g(x) | \leqq M \sum_{-\infty}^{\infty} | c_n |.$$

In view of Lemma 1, the problem may be considered as (1) the determination, if possible, of a value of a, such that the Fourier series for $\mu(t)e^{iat}$ converges absolutely, and (2) if such a value for a is determined, the evaluation of the sum of the absolute values of the Fourier coefficients or an upper bound to that sum.

Since an absolutely convergent Fourier series defines a function which, when extended periodically, is continuous, it is necessary, in order to obtain a finite bound, that the function expanded have the same value at the ends of its original interval of definition. If $| \mu(-R) | = | \mu(R) |$, θ is uniquely determined,[1] $0 \leqq \theta < 2\pi$ in the relation

$$\mu(-R) = e^{i\theta} \mu(R). \tag{1.4}$$

If

$$2aR = \theta, \tag{1.5}$$

the function $\mu(t)e^{iat}$ of Lemma 1 will satisfy the necessary condition. If $| \mu(R) | \neq | \mu(-R) |$, $\mu(t)$ can be modified, say as $\mu(t) + \alpha$, for suitable α, so the necessary condition is satisfied. We suppose throughout that (1.4) is true and that θ and a are related by (1.5).

[1] We adopt the convention that $\theta = 0$, if $\mu(R) = \mu(-R) = 0$.

Theorem 2. *If $\mu(t)$ is an absolutely continuous function satisfying* (1.4), *and if $\mu'(t)$ is equal, for almost all t, to $k(t)$, a function of bounded variation on $[-R, R]$, then*[2]

$$| g(x) | \leqq \tfrac{1}{2}RM\csc^2\tfrac{1}{2}\theta\{| k(-R) | + | k(R) | + Vk(t)\}.$$

Suppose a is defined by (1.5) and $\mu(t)e^{iat} = \sum_{-\infty}^{\infty} c_n e^{ip(n)t}$, where

$$2Rc_n = \int_{-R}^{R} \mu(t) \exp \{i[a - p(n)]t\}\, dt. \tag{2.1}$$

If we integrate (2.1) by parts and substitute the value $\theta/2R$ for a,

$$2c_n = -\frac{1}{i(\frac{1}{2}\theta - n\pi)} \int_{-R}^{R} \mu'(t) \exp\left\{i\left[\frac{\theta}{2R} - p(n)t\right]\right\} dt.$$

We next replace $\mu'(t)$ by $k(t)$ and again integrate by parts. Thus

$$2c_n = \frac{R}{(\frac{1}{2}\theta - n\pi)^2} \{k(R) \exp [i(\tfrac{1}{2}\theta - n\pi)] - k(-R) \exp [-i(\tfrac{1}{2}\theta - n\pi)]\}$$
$$- \frac{R}{(\frac{1}{2}\theta - n\pi)^2} \int_{-R}^{R} \exp\left\{i\left[\frac{\theta}{2R} - p(n)t\right]\right\} dk(t),$$

and therefore

$$| c_n | \leqq \frac{R}{2} \{| k(R) | + | k(-R) | + Vk(t)\} \frac{1}{(\frac{1}{2}\theta - n\pi)^2}. \tag{2.2}$$

We now sum, and replace $\sum_{-\infty}^{\infty} \frac{1}{(\frac{1}{2}\theta - n\pi)^2}$ by its known value,[3] $\csc^2\frac{1}{2}\theta$. Theorem 2 follows by a direct application of Lemma 1.

In particular, if $\mu(t) = it$ we get the result that $| f'(x) | \leqq MR$. This is a known generalization of S. Bernstein's classical theorem for a trigonometric polynomial [2].[4] The result for a trigonometric integral can also be obtained from another theorem of S. Bernstein [1, p. 102], dealing with certain classes of entire functions.

If $\theta = 0$, Theorem 2 fails to give a finite upper bound. This difficulty can be circumvented by a slight modification of the argument following (2.2). Suppose $N = \sup | \mu(t) |$, $(-R \leqq t \leqq R)$. Make the substitution $\theta = 0$ in (2.2) and sum, omitting the term in which $n = 0$. We obtain the result that[5]

$$\sum_{-\infty}^{\infty}{}' | c_n | \leqq (R/6)\{| k(R) | + | k(-R) | + Vk(t)\}.$$

[2] The symbol $Vk(t)$ denotes the total variation of $k(t)$ on $[-R, R]$.

[3] See T. J. I'a Bromwich, *An introduction to the theory of infinite series*, London, 1926, p. 218.

[4] The numbers in brackets refer to the bibliography.

[5] The notation $\sum_{-\infty}^{\infty}{}'$ denotes the sum in which the term with subscript zero is omitted.

Also from (2.1) we see that $| c_0 | \leqq N$. These two results, in conjunction with Lemma 1, give the following

COROLLARY. *If the hypotheses of Theorem 2 hold, and if* $\mu(-R) = \mu(R)$, *then*

$$| g(x) | \leqq M\{\tfrac{1}{8}R[| k(-R) | + | k(R) | + Vk(t)] + N\},$$

where $N = \sup | \mu(t) |$, $(-R \leqq t \leqq R)$.

We concern ourselves next with

THEOREM 3. *If* $\mu(t)$ *satisfies* (1.4), a *is defined by* (1.5) *and*

$$(3.1) \qquad (-1)^n \int_{-R}^{R} \mu(t) \exp \{i[a - p(n)t]\}\, dt \geqq 0 \qquad (n = 0, \pm 1, \pm 2, \cdots),$$

then

$$| g(x) | \leqq | \mu(R) | M.$$

In order to evaluate the sum of the absolute values of the Fourier coefficients of $\mu(t)e^{iat}$, we first consider a special $f(x)$. Suppose $f(x) = \cos Rx$. Then, by relation (1.3)

$$g(x) = \sum_{-\infty}^{\infty} c_n f[p(n) - a + x]$$

$$= \sum_{-\infty}^{\infty} c_n \cos [n\pi + (x - a)R].$$

Therefore

$$(3.3) \qquad g(a) = \sum_{-\infty}^{\infty} (-1)^n c_n.$$

Now $2R(-1)^n c_n$ is just the integral in (3.1). Hence (3.1) asserts that $(-1)^n c_n \geqq 0$, which we may express as $(-1)^n c_n = | c_n |$. Thus (3.3) states that

$$(3.4) \qquad g(a) = \sum_{-\infty}^{\infty} | c_n |.$$

Since $f(x) = \cos Rx$, the function $s(t)$ in equation (0.1) is a step function with a jump of $\frac{1}{2}$ at $-R$ and a jump of $\frac{1}{2}$ at R. Hence, we see from (0.2) that $g(x) = \frac{1}{2}\mu(-R)e^{-iRx} + \frac{1}{2}\mu(R)e^{iRx}$. Since $a = \theta/2R$, we therefore have the result that

$$(3.5) \qquad g(a) = \mu(R)e^{\frac{1}{2}i\theta}.$$

By comparison of the two values of $g(a)$ we see that $\sum_{-\infty}^{\infty} | c_n | = | \mu(R) |$. The conclusion of Theorem 3 now follows from Lemma 1.

We obtain the integral analogue of a theorem of G. Szegö [5] as a

COROLLARY. *If*

$$f(x) = \int_0^R \{\cos xt\, d\alpha(t) + \sin xt\, d\beta(t)\}, \tag{3.6}$$

$$g(x) = \int_0^R \lambda(R - t)\{\cos (xt + \gamma)\, d\alpha(t) \sin (xt + \gamma)\, d\beta(t)\}, \tag{3.7}$$

where $\lambda(t)$ *is continuous and* γ *is an arbitrary real constant, if*

$$|f(x)| \leqq M \qquad (-\infty < x < \infty), \tag{3.8}$$

and if

$$\varphi(n) = \int_0^R \lambda(t) \cos \{(\gamma + n\pi)t/R\}\, dt \geqq 0 \qquad (n = 0, \pm 1, \pm 2, \cdots), \tag{3.9}$$

then

$$|g(x)| \leqq |\lambda(0)| M. \tag{3.10}$$

We first extend the domain of definition of $\lambda(t)$ by the relation $\lambda(R - t) = -\lambda(R + t)$, $(-R \leqq t < 0)$. It is then easily verified that $f(x)$ is of the form (0.1), and $g(x)$ is of the form (0.2), where $\mu(t) = \lambda(R - t) \operatorname{sgn} t\, e^{i\gamma \operatorname{sgn} t}$. We next show that relation (3.9) implies relation (3.1). Since $|\lambda(0)| = |\mu(R)|$, this will be sufficient to prove the Corollary.

By the above definition of $\mu(t)$, we see that $\mu(R) = \lambda(0)e^{i\gamma}$ and $\mu(-R) = \lambda(0)e^{-i\gamma}$. If we write γ as $2k\pi - \delta$ where k is an integer and $0 \leqq \delta < 2\pi$, we observe that $\theta = 2\delta$ and $a = \delta/R$. Under the change of variable $u = R - t$, with the use of the relation $\gamma = 2k\pi - aR$, (3.9) is transformed into

$$(-1)^{n+2k} \int_{-R}^R \mu(t) \exp \{i[a - p(2k + n)]t\}\, dt \geqq 0 \qquad (n = 0, \pm 1, \pm 2, \cdots).$$

Since $(2k + n)$ runs through the same values as n, this is the same as (3.1).

We next alter the condition on the Fourier coefficients and obtain

THEOREM 4. *If* $\mu(t)$ *satisfies* (1.4), a *is defined by* (1.5), *and*

$$\int_{-R}^R \mu(t) \exp \{i[a - p(n)t]\}\, dt \geqq 0 \qquad (n = 0, \pm 1, \pm 2, \cdots)$$

then

$$|g(x)| \leqq |\mu(0)| M.$$

The method of proof is the same as for Theorem 3. If $s(t)$ in (0.1) is a step function with a single jump of unity at $t = 0$, then $f(x) = 1$, and $g(x) = \mu(0)$. Now by (1.3) we have $g(x) = \sum_{-\infty}^{\infty} c_n f[p(n) - a + x] = \sum_{-\infty}^{\infty} c_n$. By hypothesis $c_n \geqq 0$. Hence $c_n = |c_n|$, and $g(x) = \sum_{-\infty}^{\infty} |c_n|$. The theorem follows from the comparison of the two values for $g(x)$.

As an example of either Theorem 3 or Theorem 4 we have the following result due to Szegö [5]. If

$$g(x) = f'(x)\cos\theta + \tilde{f}'(x)\sin\theta + R\sigma(x),$$

where $\tilde{f}'(x)$ is the integral conjugate to $f'(x)$,

$$\tilde{f}'(x) = \int_{-R}^{R} |t|\, e^{ixt}\, ds(t)$$

where

$$\sigma(x) = \frac{1}{R}\int_{-R}^{R} (R - |t|)e^{ixt}\, ds(t),$$

and where θ is any real number, then $|g(x)| \leqq RM$. Here

$$g(x) = \int_{-R}^{R} \{R - |t| + it\cos\theta + |t|\sin\theta\}\, ds(t),$$

and

$$\mu(t) = R - |t| + it\cos\theta + |t|\sin\theta.$$

Thus $a = (2\theta - \pi)/2R$, and

$$\mu(t)\exp\left\{\frac{i(2\theta-\pi)t}{2R}\right\} = \frac{2R^2(1-\sin\theta)}{\pi^2}\sum_{-\infty}^{\infty}\frac{[1+(-1)^n]e^{ip(n)t}}{[(\theta/\pi) - \frac{1}{2} - n]^2}.$$

Since[6] $c_{2n} \geqq 0$ and $c_{2n+1} = 0$ $(n = 0, \pm 1, \pm 2, \cdots)$, the hypotheses of both Theorems 3 and 4 are satisfied and the conclusion follows.

We next prove

THEOREM 5. *If $\mu(t)$ is an even function which is the integral of its derivative and if $\mu'(t)$ is positive and decreasing on* $[0, R]$, *then*

$$|g(x)| \leqq \frac{2}{R}\int_0^R \mu(t)\, dt - 2\mu(0).$$

The Fourier coefficients of $\lambda(t) = \mu(t) - \mu(0)$ are given by

$$2Rc_n = \int_{-R}^{R} [\mu(t) - \mu(0)]e^{-ip(n)t}\, dt.$$

If we integrate by parts and observe that $\mu'(t)$ is odd, we see that

$$\begin{aligned} c_n &= -\frac{1}{n\pi}\int_0^R \mu'(t)\sin p(n)t\, dt \\ &= -\frac{1}{n\pi}\sum_{k=0}^{n-1}\int_{kR/n}^{(k+1)R/n} \mu'(t)\sin p(n)t\, dt. \end{aligned} \tag{5.1}$$

[6] Except where $\theta = (2k + \frac{1}{2})\pi$, whence $\mu(t) = R$ and the theorem is trivial. If, in this case, c_n were defined by continuity with respect to θ, no exception would be necessary.

If $n > 0$, then in each of the intervals $[kR/n, (k+1)R/n]$ $\sin p(n)t$ takes a full half period. Also in each interval $[kR/n, (k+1)R/n]$ the integrand has a fixed sign, positive for even k and negative for odd k. Since by hypothesis $\mu'(t)$ is a decreasing function

$$\left| \int_{kR/n}^{(k+1)R/n} \mu'(t) \sin p(n)t\, dt \right| \geqq \left| \int_{(k+1)R/n}^{(k+2)R/n} \mu'(t) \sin p(n)t\, dt \right|.$$

Therefore the sum in (5.1) has adjacent terms of opposite signs and of decreasing absolute values. Hence the sum has the sign of the first term which is positive. Thus for $n > 0$, $c_n < 0$. Since $\mu(t)$ is even, $c_{-n} = c_n$; hence $c_n < 0$, $n \neq 0$. Therefore

$$\sum_{-\infty}^{\infty}{}' \,|\, c_n \,| = -2 \sum_{n=1}^{\infty} c_n = \sum_{n=1}^{\infty} \frac{2}{n\pi} \int_0^R \mu'(t) \sin p(n)t\, dt.$$

Since the partial sums of the series $\sum_{n=1}^{\infty} n^{-1} \sin p(n)t$ are uniformly bounded and $\mu'(t)$ is integrable, we may change the order of integration and summation. Now $\sum_{n=1}^{\infty} n^{-1} \sin p(n)t = \frac{1}{2}\pi[1 - (t/R)]$; hence

$$\sum_{-\infty}^{\infty}{}' \,|\, c_n \,| = \frac{1}{R} \int_0^R \mu(t)\, dt - \mu(0). \tag{5.2}$$

It is also easily verified that

$$c_0 = |\, c_0 \,| = \frac{1}{R} \int_0^R \mu(t)\, dt - \mu(0). \tag{5.3}$$

The theorem follows immediately by combination of (5.2) and (5.3) and the use of Lemma 1.

In particular if $\mu(t) = |\, t \,|^{\alpha}$, $0 < \alpha < 1$, we get a function closely akin to the fractional derivative $f^{(\alpha)}(x)$. The result for this $\mu(t)$ is $|\, g(x) \,| \leqq (2R^{\alpha})/(\alpha + 1)$.

As an interesting corollary to this theorem we have the following result obtained by G. Sokolov [4] for trigonometric polynomials $f(x)$. If we select $\mu(t)$ as a concave polygonal curve with vertices at the points $[n, \mu(n)]$, where n is an integer, and if $\mu(0) = 0$, then $|\, g(x) \,| \leqq (M/R)\left\{\mu(R) + 2 \sum_{n=1}^{R-1} \mu(R - n)\right\}$.

Part II

We consider in this section the fractional derivative $f^{(\alpha)}(x)$, $0 < \alpha < 1$. This is defined for $f(x)$ of the type (0.1) by the equation

$$f^{(\alpha)}(x) = \int_{-R}^{R} i^{\alpha} t^{\alpha} e^{ixt}\, ds(t),$$

where $i^{\alpha} = e^{i\alpha\pi/2}$, and for negative t we understand t^{α} to be defined as $(-t)^{\alpha} e^{i\alpha\pi}$. For $f(x)$ a trigonometric polynomial, W. E. Sewell [3, p. 111] proved that

$|f^{(\alpha)}(x)| \leqq K(\alpha)MR^{\alpha}$, where $K(\alpha)$ is a function of α alone. However no explicit value was given for $K(\alpha)$. We obtain in this section a bound for $|f^{(\alpha)}(x)|$ for any $f(x)$ of the form (0.1), which is of the same order in R as Sewell's. Our result shows in particular that we may take $K(\alpha) < 7/\alpha$.

For the purpose of our proof we may consider the function $g(x) = \int_{-R}^{R} t^{\alpha} e^{ixt}\, ds(t)$, as $|g(x)| = |f^{(\alpha)}(x)|$. Throughout the proof we use the following notation:

$$a = \frac{\alpha\pi}{2R}; \qquad A(n) = \frac{\alpha R}{2n+\alpha}; \qquad B(n) = \frac{(2-\alpha)R}{(2n-\alpha)};$$

$$q(n, t) = [p(n) - a]t + \tfrac{1}{2}\alpha\pi; \tag{6.1}$$

$$d(n, t) = [-p(n) - a]t + \tfrac{1}{2}\alpha\pi.$$

We proceed as in Part I, and expand $\mu(t)e^{iat}$ in a Fourier series on $[-R, R]$. The Fourier coefficients c_n of $t^{\alpha}e^{iat}$ are given by

$$2Rc_n = \int_{-R}^{R} t^{\alpha} \exp\{[i - p(n)]t\}\, dt$$

$$= 2e^{\frac{1}{2}i\alpha\pi} \int_0^R t^{\alpha} \cos q(n, t)\, dt.$$

If we integrate by parts, we see that

$$c_n = \frac{-\alpha e^{\frac{1}{2}i\alpha\pi}}{\pi(n - \frac{1}{2}\alpha)} \int_0^R t^{\alpha-1} \sin q(n, t)\, dt.$$

For $n > 0$,

$$c_n = \frac{-\alpha e^{\frac{1}{2}i\alpha\pi}}{\pi(n - \frac{1}{2}\alpha)} \left\{\int_0^{B(n)} t^{\alpha-1} \sin q(n, t)\, dt + \int_{B(n)}^{R} t^{\alpha-1} \sin q(n, t)\, dt\right\}. \tag{6.2}$$

Since $n > 0$ and $0 < \alpha < 1$, in the first integral in (6.2) the integrand is positive; therefore the integral is positive. In the second integral of (6.2), $t^{\alpha-1}$ is decreasing. Hence, if we express the integral as a sum of integrals over complete half periods of $\sin q(n, t)$, we see that the integral is negative by applying the same argument as in the proof of Theorem 5. Hence

$$|c_n| \leqq \frac{\alpha}{\pi(n - \frac{1}{2}\alpha)} \left\{2\int_0^{B(n)} t^{\alpha-1} \sin q(n, t)\, dt - \int_0^{R} t^{\alpha-1} \sin q(n, t)\, dt\right\}. \tag{6.3}$$

Also for $n > 0$,

$$c_{-n} = \frac{\alpha e^{\frac{1}{2}i\alpha\pi}}{\pi(n + \frac{1}{2}\alpha)} \left\{\int_0^{A(n)} t^{\alpha-1} \sin d(n, t)\, dt + \int_{A(n)}^{R} t^{\alpha-1} \sin d(n, t)\, dt\right\}. \tag{6.4}$$

By the same argument as for (6.2), the first integral in (6.4) is positive and the second integral is negative. Therefore

$$(6.5)\qquad |c_{-n}| \leqq \frac{\alpha}{\pi(n+\frac{1}{2}\alpha)}\left\{2\int_0^{A(n)} t^{\alpha-1}\sin d(n,t)\,dt - \int_0^R t^{\alpha-1}\sin d(n,t)\,dt\right\}.$$

We now combine (6.3) and (6.5) and sum, obtaining

$$\begin{aligned}\sum_{n=1}^{\infty}(|c_n|+|c_{-n}|) &= \sum_{-\infty}^{\infty}{}'\,|c_n| \\ &= \frac{2\alpha}{\pi}\sum_{n=1}^{\infty}\frac{1}{(n-\frac{1}{2}\alpha)}\int_0^{B(n)} t^{\alpha-1}\sin q(n,t)\,dt \\ &\qquad + \frac{2\alpha}{\pi}\sum_{n=1}^{\infty}\frac{1}{(n+\frac{1}{2}\alpha)}\int_0^{A(n)} t^{\alpha-1}\sin d(n,t)\,dt \\ &\qquad - \frac{\alpha}{\pi}\sum_{n=1}^{\infty}\int_0^R t^{\alpha-1}\left\{\frac{\sin q(n,t)}{(n-\frac{1}{2}\alpha)}+\frac{\sin d(n,t)}{(n+\frac{1}{2}\alpha)}\right\}dt \\ &= s_1+s_2-s_3.\end{aligned}$$

Now,

$$s_1 \leqq \frac{2\alpha}{\pi}\sum_{n=1}^{\infty}\frac{1}{(n-\frac{1}{2}\alpha)}\int_0^{B(n)} t^{\alpha-1}\,dt \leqq \frac{2}{\pi}\,\frac{\alpha+2}{\alpha(2-\alpha)}\,R^{\alpha}.$$

Similarly,

$$s_2 \leqq \frac{2}{\alpha\pi}R^{\alpha}.$$

We next consider the absolute value of s_3. Replacing $q(n,t)$ and $d(n,t)$ by their values from (6.1), we have

$$\begin{aligned}|s_3| &\leqq \frac{\alpha^2}{\pi}\left|\sum_{n=1}^{\infty}\int_0^R t^{\alpha-1}\cos\left[\frac{\alpha\pi}{2}-\frac{\alpha\pi}{2R}t\right]\frac{\sin p(n)t}{(n^2-\frac{1}{4}\alpha^2)}\,dt\right| \\ &\qquad + \frac{2\alpha}{\pi}\left|\sum_{n=1}^{\infty}\int_0^R \sin\left[\frac{\alpha\pi}{2}-\frac{\alpha\pi}{2R}t\right]\frac{n\cos p(n)t}{(n^2-\frac{1}{4}\alpha^2)}\,dt\right| \\ &= s_4+s_5.\end{aligned}$$

We have

$$\begin{aligned}s_4 &\leqq \frac{\alpha^2}{\pi}\sum_{n=1}^{\infty}\frac{1}{(n^2-\frac{1}{4}\alpha^2)}\int_0^R t^{\alpha-1}\,dt \\ &\leqq \frac{\alpha}{\pi}\left\{\frac{4}{4-\alpha^2}+\frac{1}{\alpha}\log\frac{2-\alpha}{2+\alpha}\right\}R^{\alpha}.\end{aligned}$$

Moreover, for $N = 1, 2, \cdots$,

$$\left|t^{\alpha-1}\sum_{n=1}^{N}\frac{n\cos p(n)t}{(n^2-\frac{1}{4}\alpha^2)}\right| \leqq C(\alpha)t^{\alpha-1} + t^{\alpha-1}\left|\sum_{n=1}^{N}\frac{\cos p(n)t}{n}\right|,$$

where $C(\alpha)$ depends only on α. It is easily verified, say by partial summation, that the sums $\left| \sum_{n=1}^{N} n^{-1} \cos p(n)t \right|$ are dominated by an integrable function. Hence we may change the order of integration and summation in s_5. Thus

$$s_5 \leqq \frac{\alpha^3}{2\pi} \int_0^R t^{\alpha-1} \left| \sum_{n=1}^{\infty} \frac{\cos p(n)t}{n(n^2 - \frac{1}{4}\alpha^2)} \right| dt + \frac{2\alpha}{\pi} \int_0^R \left| \sum_{n=1}^{\infty} \frac{\cos p(n)t}{n} \right| dt$$

$$\leqq \left\{ \frac{\alpha^2}{\pi} \left[\frac{2}{4 - \alpha^2} + \frac{1}{\alpha^2} \log \frac{1}{4 - \alpha^2} \right] + \frac{4 \log 2}{\pi} \right\} R^\alpha - \frac{2\alpha}{\pi} \int_0^R t^{\alpha-1} \log \sin \frac{\pi}{2R} t\, dt$$

$$\leqq \left\{ \frac{\alpha^2}{\pi} \left[\frac{2}{4 - \alpha^2} + \frac{1}{\alpha^2} \log \frac{1}{4 - \alpha^2} \right] + \frac{4 \log 2}{\pi} + \frac{1}{\alpha} \right\} R^\alpha.$$

We also have

$$| c_0 | \leqq \frac{1}{2R} \int_{-R}^{R} | \mu(t) | \, dt = \frac{R^\alpha}{\alpha + 1}.$$

If we combine the bounds for s_1, s_2, s_4, s_5 with that for $| c_0 |$, we see that

$$\sum_{-\infty}^{\infty} | c_n | \leqq \left\{ \frac{8 + 2\alpha}{\alpha\pi(2 - \alpha)} + \frac{2\alpha + 1}{\alpha(1 + \alpha)} + \frac{4}{\pi} \log 2 - \log (2 + \alpha) \right\} R^\alpha \leqq \frac{7R^\alpha}{\alpha}.$$

Therefore by application of Lemma 1, we have

THEOREM 6. *If $f(x)$ is of the form (0.1) and $f^{(\alpha)}(x)$, $0 < \alpha < 1$, is its fractional derivative, then*

$$|f^{(\alpha)}(x)| \leqq \left\{ \frac{8 + 2\alpha}{\alpha\pi(2 - \alpha)} + \frac{2\alpha + 1}{\alpha(1 + \alpha)} + \frac{4}{\pi} \log 2 - \log (2 + \alpha) \right\} MR^\alpha.$$

BIBLIOGRAPHY

1. S. BERNSTEIN, *Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle*, Paris, 1926.
2. R. P. BOAS, JR., *The derivative of a trigonometric integral*, Journal of the London Mathematical Society, vol. 12(1937), pp. 164–165.
3. W. E. SEWELL, *Generalized derivatives and approximation by polynomials*, Transactions of the American Mathematical Society, vol. 41(1937), pp. 84–123.
4. G. T. SOKOLOV, *Sur quelques propriétés extrémales des sommes trigonométriques*, Bulletin de l'Académie des Sciences de l'URSS (Izvestiya Akademii Nauk SSSR), ser. 7, Classe des sciences mathématiques et naturelles, 1935, pp. 857–884.
5. G. SZEGÖ, *Über ein Satz der Herrn Serge Bernstein*, Schriften der Königsberger Gelehrten Gesellschaft, Naturwissenschaftliche Klasse, vol. 5(1928), pp. 59–70.

DUKE UNIVERSITY.

THE FUNCTION OF MEAN CONCENTRATION OF A CHANCE VARIABLE

By Tatsuo Kawata

1. Introduction

1.1. Let X be a one-dimensional chance variable which is defined by its probability distribution function

$$Pr(X < x) = \sigma(x).$$

Thus $\sigma(x)$ is a non-decreasing function such that $\sigma(-\infty) = 0$ and $\sigma(\infty) = 1$. Let $\{X_n\}$ be a sequence of independent chance variables; that is, let the k-dimensional chance variable $(X_{i_1}, X_{i_2}, \cdots, X_{i_k})$ be defined by the condition

$$Pr(X_{i_1} < x_1, X_{i_2} < x_2, \cdots, X_{i_k} < x_k) = Pr(X_{i_1} < x_1)Pr(X_{i_2} < x_2) \cdots Pr(X_{i_k} < x_k),$$

for every finite set of distinct integers $i_1, i_2, \cdots, i_k$ and for every set of real numbers $x_1, x_2, \cdots, x_k$.

Consider the series of independent chance variables

$$\sum_{n=1}^{\infty} X_n. \tag{1.1}$$

The series (1.1) is said to converge in probability if

$$Pr(| S_n - S | > \epsilon) \to 0 \qquad \text{as } n \to \infty,$$

for every $\epsilon > 0$ for some chance variable S, where S_n denotes the partial sum $X_1 + X_2 + \cdots + X_n$. The convergence problem of (1.1) was treated by a great number of writers.

Among many results concerning the convergence problem of (1.1), there are two theories, one of which is due to A. Khintchine and A. Kolmogoroff ([11]; see also [5], [8], [12], [13] and [15], p. 142)[1] and the other due to P. Lévy ([14]; [15], pp. 130–140). A main theorem in the former theory is the one which gives the necessary and sufficient conditions for the convergence in probability of (1.1) in terms of expectations of X_i and X_i^2 under certain hypotheses. The central idea in Lévy theory is to use the function of maximum concentration.

Let the distribution function of a chance variable X be $\sigma(x)$. The function

$$Q(h) = \max_{-\infty<x<\infty} \{\sigma(x + h + 0) - \sigma(x - h - 0)\} \tag{1.2}$$

Received May 25, 1940; in revised form, April 30, 1941.

[1] Numbers in brackets refer to the bibliography at the end.

is following Lévy called the function of maximum concentration of X ([14]; [15], p. 44). Let the function of maximum concentration of a partial sum

$$S_{n,N} = \sum_{i=n+1}^{N} X_i$$

be $Q_{n,N}(h)$. It is known that the function of maximum concentration never increases if a chance variable is added ([15], p. 90). Thus $Q_{n,N}(h)$ does not increase as N increases and does not decrease as n increases. Hence

$$Q_n(h) = \lim_{N\to\infty} Q_{n,N}(h), \qquad Q(h) = \lim_{n\to\infty} Q_n(h) \tag{1.3}$$

are well defined. The following theorem is one of the fundamental theorems in the Lévy theory.

THEOREM A. *The function* $Q(h)$ *is independent of* h *and is either* 0 *or* 1. *If* $Q(h) = 1$, *then there exists a sequence of numbers* $\{a_n\}$ *such that* $S_n - a_n$ *converges in distribution; that is, the distribution function of* $S_n - a_n$ *converges to some distribution function and if* $Q(h) = 0$, *then* $S_n - a_n$ *cannot converge in distribution for any number sequence* $\{a_n\}$ ([14]; [15], p. 130).

1.2. Let $\sigma(x)$ be the distribution function of a chance variable X and consider the function

$$f(t) = \int_{-\infty}^{\infty} e^{itx}\, d\sigma(x) \tag{1.4}$$

which is evidently continuous and is such that $f(0) = 1, |f(t)| \leqq 1$. This function is called the Fourier-Stieltjes transform (characteristic function) of X ([15], p. 37). Let X_1 and X_2 be mutually independent chance variables and let their Fourier-Stieltjes transforms be $f_1(t)$ and $f_2(t)$, respectively. Then the characteristic function of $X = X_1 + X_2$ is represented as the product $f_1(t)f_2(t)$. The convergence in probability of S_n, the partial sum of (1.1), is equivalent to the uniform convergence in every finite interval of $\prod_{i=1}^{n} f_i(t)$, where $f_i(t)$ is the characteristic function of X ([15], p. 48).

Now for convenience we assume throughout this paper that a distribution function $\sigma(x)$ is normalized as

$$\sigma(x) = \tfrac{1}{2}\{\sigma(x + 0) + \sigma(x - 0)\}.$$

Then the distribution function is by the Lévy inversion formula given by its Fourier-Stieltjes transform as

$$\sigma(x) - \sigma(0) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{1 - e^{-itx}}{it} f(t)\, dt.$$

We obtain from this that

$$\sigma(x + h) - \sigma(x - h) = \frac{1}{\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{\sin ht}{t} e^{-itx} f(t)\, dt.$$

Since for every $h > 0$ we have[2] $\sin ht\, f(t)/t \in L_2(-\infty, \infty)$, $(\frac{1}{2}\pi)^{\frac{1}{2}}\{\sigma(x + h) - \sigma(x - h)\}$ is the ordinary Fourier transform of the function $\sin ht\, f(t)/t$. Thus Plancherel's theorem shows that

$$\int_{-\infty}^{\infty} \{\sigma(x + h) - \sigma(x - h)\}^2 dx = \frac{2}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2 ht}{t^2} \mid f(t) \mid^2 dt. \tag{1.5}$$

Since $f(-t) = \overline{f(t)}$ by definition, $| f(t) |^2$ is an even function of t. Hence (1.5) becomes

$$\frac{1}{2h} \int_{-\infty}^{\infty} \{\sigma(x + h) - \sigma(x - h)\}^2 dx = \frac{2}{\pi} \int_{0}^{\infty} \frac{\sin^2 ht}{ht^2} \mid f(t) \mid^2 dt. \tag{1.6}$$

We now call the left side of (1.6) *the function of mean concentration of* X and denote as $C(h)$. In this paper we make use of this function instead of the function of maximum concentration, and we prove some theorems in the Lévy theory by Fourier analysis. This will be done in Part 3. In Part 4 the summability problem, which has been also treated by Lévy [16], shall be considered by the similar method. Lastly, we mention that the behavior of $C(h)$ at $h = 0$ was discussed by N. Wiener ([23], [24]) in connection with his spectrum theory of a function.

2. Properties of the function of mean concentration

2.1. THEOREM 1.

$$\lim_{h \to \infty} C(h) = 1, \tag{2.1}$$

$$\lim_{h \to 0} C(h) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \mid f(t) \mid^2 dt. \tag{2.2}$$

This is known (see, e.g., [2] and [23], [24]). (2.2) is due to N. Wiener ([23], [24]).

THEOREM 2. *$C(h)$ is a non-decreasing function of h (> 0).*

Let

$$\tau(x) = \int_{-\infty}^{\infty} \sigma(x - t)\, d(1 - \sigma(-t)).$$

This is a symmetric distribution function in the sense that

$$\tau(x) + \tau(-x) = 1,$$

and the Fourier-Stieltjes transform is, as is readily verified, $| f(t) |^2$. Thus the Lévy inversion formula shows that

$$\tau(x) - \frac{1}{2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1 - e^{-itx}}{it} \mid f(t) \mid^2 dt$$

$$= \frac{1}{\pi} \int_{0}^{\infty} \frac{\sin xt}{t} \mid f(t) \mid^2 dt.$$

[2] $f(t) \in L_2$ means that $f(t)$ is measurable and $\int_{-\infty}^{\infty} \mid f(t) \mid^2 dt < \infty$.

Since $|f(t)|^2/t \in L_2(1, \infty)$, we have

$$\tau(x) - \frac{1}{2} = \frac{1}{\pi}\int_0^1 \frac{\sin xt}{t}\,|f(t)|^2\,dt + \underset{T\to\infty}{\text{l.i.m.}}\,\frac{1}{\pi}\int_1^T \frac{\sin xt}{t}\,|f(t)|^2\,dt,$$

where l.i.m. means the limit in the mean in $L_2(0, \infty)$. Thus a known property of weak convergence shows that

$$\begin{aligned}\int_0^{2h} \{\tau(x) - \tfrac{1}{2}\}\,dx &= \frac{1}{\pi}\int_0^{2h} dx \int_0^1 \frac{\sin xt}{t}\,|f(t)|^2\,dt \\ &\qquad + \int_0^{2h} dx \,\frac{1}{\pi}\, \underset{T\to\infty}{\text{l.i.m.}} \int_1^T \frac{\sin xt}{t}\,|f(t)|^2\,dt \\ &= \frac{1}{\pi}\int_0^1 |f(t)|^2\,dt \int_0^{2h} \frac{\sin xt}{t}\,dx + \lim_{T\to\infty}\frac{1}{\pi}\int_1^T \frac{|f(t)|^2}{t}\,dt\int_0^{2h}\sin xt\,dx \\ &= \frac{2}{\pi}\int_0^\infty \frac{\sin^2 ht}{t^2}\,|f(t)|^2\,dt.\end{aligned}$$

Hence we get

$$C(h) = \frac{1}{h}\int_0^{2h} \{\tau(x) - \tfrac{1}{2}\}\,dx. \tag{2.3}$$

Since $\Phi(x) = \tau(x) - \frac{1}{2}$ is non-decreasing, $\Phi(+0) = 0$, $\Phi(\infty) = \frac{1}{2}$, and $C(h)$ is also non-decreasing. For $h > h'$

$$\begin{aligned}C(h) - C(h') &= \frac{1}{h}\int_0^{2h} \Phi(x)\,dx - \frac{1}{h'}\int_0^{2h'} \Phi(x)\,dx \\ &= \frac{1}{h}\int_{2h'}^{2h} \Phi(x)\,dx - \frac{h-h'}{hh'}\int_0^{2h'} \Phi(x)\,dx \\ &\geqq 2\,\frac{h-h'}{h}\,\Phi(2h') - 2\,\frac{h-h'}{h}\,\Phi(2h') = 0.\end{aligned}$$

2.2. Let

$$\epsilon(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0, \end{cases}$$

which is called the unit distribution function. We give here a criterion in terms of $C(h)$ in order that a distribution should be the unit distribution function. We use the following lemma which is essentially known (cf. [25]).

LEMMA 1. *If $f(t)$ is the Fourier-Stieltjes transform of a distribution function $\sigma(x)$ and $|f(t_1)| = |f(t_2)| = 1$, where t_1/t_2 is irrational, then $\sigma(x) = \epsilon(x - a)$ for some a, $\epsilon(x)$ being the unit distribution function.*

If $|f(t_1)| = 1$, then for some constant w,

$$\int_{-\infty}^{\infty} e^{it_1x+iw}\,d\sigma(x) = 1$$

and thus

$$\int_{-\infty}^{\infty} \{1 - \cos(t_1x + w)\}\,d\sigma(x) = 0.$$

Hence $\sigma(x)$ is constant except possibly at the values of x such that $t_1x + w = 2k\pi$ ($k = 0, \pm 1, \cdots$). Thus the points of actual increase of $\sigma(x)$ are contained in an arithmetical progression with a common difference $2\pi/t_1$; that is, the distances of points of increase are integral multiples of $2\pi/t_1$. If further $|f(t_2)| = 1$, then the distances of points of increase are also integral multiples of $2\pi/t_2$. Therefore $\sigma(x)$ has a unique point of actual increase.

THEOREM 3. *$\sigma(x)$ is $\epsilon(x - a)$ for some a, $\epsilon(x)$ being the unit distribution, if and only if $C(h) = 1$ for every $h > 0$.*

If $\sigma(x) = \epsilon(x - a)$, then $f(t) = e^{ita}$, thus $|f(t)| = 1$. Hence

$$C(h) = \frac{2}{\pi}\int_0^\infty \frac{\sin^2 ht}{ht^2}\,dt = 1,$$

for every $h > 0$. Conversely, if $C(h) = 1$ for every $h > 0$, then

$$\frac{2}{\pi}\int_0^\infty \frac{\sin^2 ht}{ht^2}(1 - |f(t)|^2)\,dt = 0.$$

Since $|f(t)| \leqq 1$, $|f(t)|^2 = 1$ almost everywhere. Since $f(t)$ is continuous, $|f(t)|^2 = 1$ everywhere. Thus $|f(t)| = 1$. Lemma 1 shows our assertion.

THEOREM 4. *Let the Fourier-Stieltjes transforms of distribution functions $\sigma_1(x)$ and $\sigma_2(x)$ be $f_1(t)$ and $f_2(t)$ respectively and the functions of mean concentration be $C_1(h)$ and $C_2(h)$ respectively. If*

$$|f_1(t)| \leqq |f_2(t)|, \tag{2.4}$$

then

$$C_1(h) \leqq C_2(h). \tag{2.5}$$

This is immediate from the definition (1.6).

2.3. Next we consider here the relation between the maximum concentration $Q(h)$ and the mean concentration.

THEOREM 5. *If the function of maximum concentration and the function of mean concentration of a distribution function $\sigma(x)$ be $Q(h)$ and $C(h)$ respectively, then*

$$Q(2h) \geqq C(h) \geqq \tfrac{1}{2}Q^2(h). \tag{2.6}$$

For

$$\begin{aligned} C(h) &= \frac{1}{2h}\int_{-\infty}^{\infty} \{\sigma(x + h) - \sigma(x - h)\}^2\,dx \\ &\leqq Q(2h)\frac{1}{2h}\int_{-\infty}^{\infty} \{\sigma(x + h) - \sigma(x - h)\}\,dx \\ &= Q(2h)\lim_{T\to\infty}\left\{\frac{1}{2h}\int_{T-h}^{T+h} \sigma(x)\,dx + \frac{1}{2h}\int_{-T-h}^{-T+h} \sigma(x)\,dx\right\} = Q(2h), \end{aligned}$$

and this proves the left part of (2.6).

Next we take ξ such that

$$Q(h) = \max_{-\infty<x<\infty} \{\sigma(x + h + 0) - \sigma(x - 0)\}$$
$$= \sigma(\xi + \tfrac{1}{2}h + 0) - \sigma(\xi - \tfrac{1}{2}h - 0).$$

Then clearly in the interval $\xi - \frac{1}{2}h < x < \xi + \frac{1}{2}h$,

$$\sigma(x + h) - \sigma(x - h) \geqq Q(h).$$

Thus

$$C(h) = \frac{1}{2h}\int_{-\infty}^{\infty} \{\sigma(x + h) - \sigma(x - h)\}^2\,dx$$
$$\geqq \frac{1}{2h}\int_{\xi-\frac{1}{2}h}^{\xi+\frac{1}{2}h} \{\sigma(x + h) - \sigma(x - h)\}^2\,dx$$
$$\geqq \tfrac{1}{2}Q^2(h),$$

and the proof of (2.6) is complete.

3. Sums of independent chance variables

3.1. Let X_1 and X_2 be independent chance variables and let their distribution functions be $\sigma_1(x)$ and $\sigma_2(x)$, respectively. Let the functions of mean concentration of $\sigma_1(x)$ and $\sigma_2(x)$ be $C_1(h)$ and $C_2(h)$ respectively. Further we let the distribution and the function of mean concentration of a chance variable $Y = X_1 + X_2$ be $\sigma(x)$ and $C(h)$, respectively.

THEOREM 6. *For every $h > 0$, we have*

$$C(h) \leqq C_i(h) \qquad (i = 1, 2). \tag{3.1}$$

If

$$C_2(h) = C(h), \qquad \textit{for some } h > 0, \tag{3.2}$$

then $\sigma_1(x) = \epsilon(x - a)$ for some a, where $\epsilon(x)$ is the unit distribution function.

By the well-known fact, the Fourier-Stieltjes transform of $\sigma(x)$ is $f_1(t)f_2(t)$, where $f_i(t)$ is the Fourier-Stieltjes transform of $X_i(t)$ $(i = 1, 2)$. The first part of the theorem is trivial.

If (3.2) holds, then

$$\int_0^{\infty} \frac{\sin^2 ht}{ht^2}\{|f_2(t)|^2 - |f_1(t)f_2(t)|^2\}\,dt = 0.$$

Hence $|f_2(t)|^2(1 - |f_1(t)|^2) = 0$ everywhere. If $|f_1(t)|^2 = 1$ for every t, then by Lemma 1, $\sigma_1(x) = \epsilon(x - a)$. And if there exists a set of t-values such that $|f_2(t)| = 0$, this set is closed since $f_2(t)$ is continuous. Thus in its contiguous interval $|f_1(t)| = 1$ and the assertion follows by Lemma 1. Since $|f_2(t)| = 0$ cannot hold everywhere, the conclusion is proved.

THEOREM 7. *Let $X_1, X_2, \cdots$ be a sequence of independent chance variables and put*

$$S_n = X_1 + X_2 + \cdots + X_n. \tag{3.3}$$

Then $C_n(h)$ the function of mean concentration of S_n is non-increasing with respect to n for every $h > 0$.

This is immediate.

3.2. Put

$$S_{n,m} = X_{n+1} + X_{n+2} + \cdots + X_m, \tag{3.4}$$

and let its function of mean concentration be $C_{n,m}(h)$. Then by Theorem 7, $C_{n,m}(h)$ is non-increasing with respect to m for every h. Thus $\lim_{m\to\infty} C_{n,m}(h) = C_n(h)$ exists. It is also evident that $C_n(h)$ is non-decreasing with respect to n for every h. Thus $\lim_{n\to\infty} C_n(h) = C(h)$ exists and is called the function of limit mean concentration of $\{X_n\}$. Now we shall prove

THEOREM 8. *$C(h)$ the function of the limit mean concentration of a sequence of independent chance variables $\{X_n\}$ is either 0 for every $h > 0$ or 1 for every $h > 0$.*

For the proof we require a simple lemma.

LEMMA 2. *Let $\varphi(t)/(1 + t^2) \in L_1(0, \infty)$. Then if*

$$\int_0^\infty \frac{\sin^2 xt}{t^2}\varphi(t)\,dt = 0 \tag{3.5}$$

holds for every x, $\varphi(t) = 0$ for almost all values of t.

Let ϵ be any positive number. Then (3.5) yields that

$$\int_0^\infty \sin^2 \frac{x+\epsilon}{2}t\cdot\frac{\varphi(t)}{t^2}\,dt = 0,$$

$$\int_0^\infty \sin^2 \frac{x-\epsilon}{2}t\cdot\frac{\varphi(t)}{t^2}\,dt = 0.$$

The subtraction shows that

$$\int_0^\infty \sin xt\,\frac{\sin \epsilon t}{t^2}\varphi(t)\,dt = 0,$$

for every x. Putting in this equation $x + \epsilon$ and $x - \epsilon$ for x and subtracting we obtain

$$\int_0^\infty \cos xt\,\frac{\sin^2 \epsilon t}{t^2}\varphi(t)\,dt = 0.$$

The unicity theorem on cosine transform shows that

$$\frac{\sin^2 \epsilon t}{t^2}\varphi(t) = 0$$

holds at almost all values of t. Thus $\varphi(t) = 0$ almost everywhere.

Now we prove Theorem 8. The Fourier-Stieltjes transform of the distribution function of $S_{n,m}$ is, as is well known, $f_{n+1}(t) \cdots f_m(t)$. If we put

$$\alpha_{n,m}(t) = | f_{n+1}(t) \cdots f_m(t) |^2,$$

then clearly $\lim_{m\to\infty} \alpha_{n,m}(t)$ exists at every t, since $\alpha_{n,m}(t)$ is non-increasing. Thus by the well-known fact on infinite products

$$\alpha(t) = \lim_{n\to\infty} \lim_{m\to\infty} \alpha_{n,m}(t) \tag{3.6}$$

is either zero or 1 at each value of t. Further we have

$$\begin{aligned} C(h) &= \lim_{n\to\infty} \lim_{m\to\infty} \frac{2}{\pi} \int_0^\infty \frac{\sin^2 ht}{ht^2} \alpha_{n,m}(t)\, dt \\ &= \frac{2}{\pi} \int_0^\infty \frac{\sin^2 ht}{ht^2} \alpha(t)\, dt. \end{aligned} \tag{3.7}$$

Now by (2.3) we can write

$$C(h) = \lim_{n\to\infty} \lim_{m\to\infty} \frac{1}{h} \int_0^{2h} \{\tau_{n,m}(x) - \tfrac{1}{2}\}\, dx,$$

where the definition of $\tau_{n,m}(x)$ will be evident. Since $\tau_{n,m}(x)$ is a non-decreasing function of x and $\tau_{n,m}(+0) = \frac{1}{2}$, $\tau_{n,m}(\infty) = 1$, by repeated applications of the Helly theorem and the diagonal method we can prove that there exists a non-decreasing function $\tau(x)$ such that

$$C(h) = \frac{1}{h} \int_0^{2h} \{\tau(x) - \tfrac{1}{2}\}\, dx,$$

where $\frac{1}{2} \leqq \tau(x) \leqq 1$, for $x > 0$.

From (3.7) we get

$$\frac{1}{h} \int_0^{2h} \{\tau(x) - \tfrac{1}{2}\}\, dx = \frac{2}{\pi} \int_0^\infty \frac{\sin^2 ht}{ht^2} \alpha(t)\, dt = C(h). \tag{3.8}$$

If $\tau(\infty) = \frac{1}{2}$, then $\tau(x) = \frac{1}{2}$ at all points x. Thus by (3.8), $C(h) = 0$ for every $h > 0$. Next let $\tau(\infty) = \beta \neq \frac{1}{2}$, $\frac{1}{2} < \beta \leqq 1$. And we define a symmetric distribution function $v(x)$ as follows:

$$\begin{aligned} v(x) &= \tau(x)/\beta, && \text{for } x > 0, \\ v(x) &= 1 - \tau(-x)/\beta, && \text{for } x < 0. \end{aligned}$$

Let the Fourier-Stieltjes transform of $v(x)$ be $\chi(t)$. Then the Lévy inversion formula shows that

$$v(x) - \frac{1}{2} = \frac{1}{\pi} \int_0^\infty \frac{\sin xt}{t} \chi(t)\, dt. \tag{3.9}$$

By integrating both sides over $(0, 2h)$ and dividing by $2h$, we obtain

$$\frac{1}{2h}\int_0^{2h} v(x)\,dx - \frac{1}{2} = \frac{1}{\pi}\int_0^{\infty} \frac{\sin^2 ht}{ht^2}\,\chi(t)\,dt,$$

or we have

$$\frac{1}{2h\beta}\int_0^{2h} \tau(x)\,dx - \frac{1}{2} = \frac{1}{\pi}\int_0^{\infty} \frac{\sin^2 ht}{ht^2}\,\chi(t)\,dt.$$

Comparing with (3.8), we get

$$\frac{\beta}{2} + \frac{\beta}{\pi}\int_0^{\infty} \frac{\sin^2 ht}{ht^2}\,\chi(t)\,dt = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \frac{\sin^2 ht}{ht^2}\,\alpha(t)\,dt.$$

Thus

$$\frac{1}{\pi}\int_0^{\infty} \frac{\sin^2 ht}{ht^2}\,\{\alpha(t) + 1 - \beta - \beta\chi(t)\}\,dt = 0,$$

and it results, by Lemma 2, that

$$\frac{\{\alpha(t) + 1 - \beta\}}{\beta} = \chi(t),$$

almost everywhere. Since $\alpha(t)$ is zero or 1 at each value t, $\chi(t)$ is either $1/\beta - 1$ or $2/\beta - 1$ at each t. Since $\chi(t)$ is continuous, it is either $1/\beta - 1$ for every t or $2/\beta - 1$ for every t. Thus $\chi(t)$ is a constant which is evidently 1 since $\chi(0) = 1$ and $\beta = 1$. Hence by (3.9) $v(x) = 1$ everywhere, and this means that $\tau(x) = 1$ for $x > 0$. Thus by (3.8) $C(h) = 1$ for every $h > 0$. Thus the theorem is proved.

We notice that from the above theorem it is evident that the limit maximum concentration $Q(h)$ is either 0 or 1 for every h. This is Lévy's conclusion.

4. Summability of a series of independent chance variables

As in the preceding part, let $\{X_i\}$ be a sequence of independent chance variables and put

$$S_n = \sum_{i=1}^{n} X_i.$$

Let $a_{i,k}$ $(i, k = 1, 2, \cdots)$ be non-negative numbers such that

$$\sum_{i=1}^{\infty} a_{i,k} = 1, \tag{4.1}$$

$$\lim_{k\to\infty} a_{i,k} = 0. \tag{4.2}$$

Further let

$$\sigma_k = \sum_{i=1}^{\infty} a_{i,k} S_i, \tag{4.3}$$

the convergence (almost certainly) of the right side being supposed. If the distribution of σ_k converges to some distribution, we say that the series $\sum X_i$ is summable in distribution. Summability in probability is similarly defined.

THEOREM 9. *If $\sum (X_i - a_i)$ does not converge in distribution for any sequence $\{a_n\}$, then $\sum (X_i - a_i)$ is not summable in distribution (consequently not summable in probability) for any sequence $\{a_n\}$.*

This theorem is also due to P. Lévy [16].

LEMMA 3. *If X_i converges in distribution to X, then the function of mean concentration of X_i tends to that of X everywhere.*

If X_i tends to X in distribution, then the Fourier-Stieltjes transform of X tends to that of X and the integrand of

$$\int_0^\infty \frac{\sin^2 ht}{ht^2} \, |f_i(t)|^2 \, dt$$

is uniformly majorized by the integrable function $\sin^2 ht/ht^2$. Hence the conclusion is immediate.

We now prove the theorem. By Theorem 8 and Lévy's Theorem A we have $C(h) = \lim_{n\to\infty} \lim_{m\to\infty} C_{n,m}(h) = 0$ for every $h > 0$, using the notation of Part 3, $C_{n,m}(h)$ is the mean concentration of $\sum_{i=n+1}^{m} X_i$. We have by (4.3) and (4.1),

$$\begin{aligned}\sigma_k &= \left(S_n - \sum_{i=1}^{n} a_{i,k} S_{i,n}\right) + \left(\sum_{i=n+1}^{m} a_{i,k} S_{n,i} + S_{n,m} - \sum_{i=1}^{m} a_{i,k} S_{n,m}\right) \\ &\qquad + \sum_{i=m+1}^{\infty} a_{i,k}(S_{n,i} - S_{n,m}) \quad (m > n) \\ &= U_n + T_{n,m} + V_{n,m},\end{aligned} \tag{4.5}$$

say. U_n does not contain X_i $(i > n + 1)$ and $T_{n,m}$ contains only X_i $(n + 1 \leqq i \leqq m)$ and $V_{n,m}$ does not contain X_i $(i \leqq m)$. Hence U_n, $V_{n,m}$ and $T_{n,m}$ are independent. By Theorem 7 we have

$$C(h, \sigma_k) \leqq C(h, T_{n,m}), \tag{4.6}$$

where $C(h, \sigma_k)$ and $C(h, T_{n,m})$ denote the mean concentrations of σ_k and $T_{n,m}$, respectively. Since by (4.2) $T_{n,m}$ converges to $S_{n,m}$, Lemma 3 shows that

$$\lim_{k\to\infty} C(h, T_{n,m}) = C_{n,m}(h),$$

so that by (4.6) we get

$$\overline{\lim_{k\to\infty}}\, C(h, \sigma_k) \leqq C_{n,m}(h)$$

which tends to 0 as $m \to \infty$, $n \to \infty$. Thus

$$\lim_{k\to\infty} C(h, \sigma_k) = 0, \tag{4.7}$$

for every $h > 0$, and this means that

$$\lim_{k\to\infty} \int_0^\infty \frac{\sin^2 ht}{ht^2} |f_k(t)|^2 dt = 0, \tag{4.8}$$

where f_k is the Fourier-Stieltjes transform. If, for some sequence $\{a_n\}$, $\sum (X_i - a_i)$ is summable, then there exists a sequence of numbers $\{b_n\}$ such that $\sigma_k - b_k$ converges in distribution. This means that $e^{ib_kt} | f_k(t) |$ must converge uniformly in every finite interval to the Fourier-Stieltjes transform of a distribution, and (4.8) is contradicted.

Bibliography

1. S. Bochner, *Vorlesungen über Fouriersche Integrale*, Leipzig, 1932.
2. H. Cramér, *Random Variables and Probability Distributions*, Cambridge Tracts, no. 36, 1937.
3. B. Gnedenko, *Über die Konvergenz der Verteilungsgesetze von Summen von einander unabhängiger Summanden*, Comptes Rendus de l'Académie des Sciences de l'U.R.S.S., vol. 18(1938), pp. 231–234.
4. B. Jessen, *The theory of integration in a space of an infinite number of dimensions*, Acta Mathematica, vol. 63(1934), pp. 249–323.
5. B. Jessen and A. Wintner, *Distribution functions and the Riemann zeta-function*, Transactions of the American Mathematical Society, vol. 38(1935), pp. 48–88.
6. M. Kac, *Sur les fonctions indépendantes*, I (*Propriétés générales*), Studia Mathematica, vol. 6(1936), pp. 46–58.
7. M. Kac and H. Steinhaus, *Sur les fonctions indépendantes*, II (*La loi exponentielle, la divergence de série*), Studia Mathematica, vol. 6(1936), pp. 59–66.
8. E. R. van Kampen, *Infinite product measures and infinite convolutions*, American Journal of Mathematics, vol. 62(1939), pp. 417–448.
9. E. R. van Kampen and A. Wintner, *On divergent infinite convolutions*, American Journal of Mathematics, vol. 59 (1937), pp. 635–654.
10. A. Khintchine, *Über Klassenkonvergenz von Verteilungsgesetzen*, Bulletin de l'Institut Mathématique et Mécanique, Université de Tomsk, vol. 1(1937), pp. 258–261.
11. A. Khintchine and A. Kolmogoroff, *Über Konvergenz von Reihen, deren Glieder durch den Zufall bestimmt werden*, Recueil Mathématique, vol. 32(1925), pp. 668–677.
12. A. Kolmogoroff, *Über die Summen durch den Zufall bestimmter unabhängiger Grössen*, Mathematische Annalen, vol. 99(1928), pp. 309–319.
13. A. Kolmogoroff, *Bemerkungen zu meiner Arbeit "Über die Summen zufälliger Grössen"*, Mathematische Annalen, vol. 102(1930), pp. 484–488.
14. P. Lévy, *Sur les séries dont les termes sont des variables éventuelles indépendantes*, Studia Mathematica, vol. 3(1931), pp. 119–155.
15. P. Lévy, *Théorie de l'Addition des Variables Aléatoires*, Paris, 1937.
16. P. Lévy, *Sur la sommabilité des séries aléatoires divergentes*, Bulletin de la Société Mathématique de France, vol. 63(1935), pp. 1–35.
17. J. Marcinkiewicz and A. Zygmund, *Quelques théorèmes sur les fonctions indépendantes*, Studia Mathematica, vol. 7(1938), pp. 104–120.
18. J. Marcinkiewicz and A. Zygmund, *Sur les fonctions indépendantes*, Fundamenta Mathematicae, vol. 29(1937), pp. 60–90.
19. R. E. A. C. Paley and A. Zygmund, *On some series of functions*, I, Proceedings of the Cambridge Philosophical Society, vol. 26(1930), pp. 337–357; II, vol. 26(1930), pp. 458–474; III, vol. 28(1932), pp. 190–205.
20. H. Rademacher, *Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen*, Mathematische Annalen, vol. 87(1922), pp. 112–138.

21. H. Steinhaus, *Sur la probabilité de la convergence de séries*, Studia Mathematica, vol. 2(1930), pp. 21–39.
22. E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, Oxford, 1937.
23. N. Wiener, *The Fourier Integral and Certain of Its Applications*, Cambridge, 1933.
24. N. Wiener, *Generalized harmonic analysis*, Acta Mathematica, vol. 55(1930), pp. 117–258.
25. A. Wintner, *On a class of Fourier transforms*, American Journal of Mathematics, vol. 58(1936), pp. 45–90.
26. A. Wintner, *Gaussian distributions and convergent infinite convolutions*, American Journal of Mathematics, vol. 57(1935), pp. 821–826.

Tôhoku Imperial University.

THE SINGULARITIES OF CAUCHY'S DISTRIBUTIONS

By Aurel Wintner

Generalizing the symmetric normal distribution law ($\lambda = 2$), Cauchy[1] proposed the symmetric "stable" distributions, those having the Fourier transform $e^{-|t|^\lambda}$, where the positive index λ is unspecified. The object of the present note is the determination of the behavior of the analytic continuations of the resulting densities of probability. If $\lambda \geqq 1$, the situation is of a trivial nature. If $\lambda < 1$, the only case known today seems to be $\lambda = \frac{1}{2}$; a case which is quite accidental, since it can be reduced to the normal case $\lambda = 2$, if the Green function belonging to a certain problem concerning the parabolic equation $u_x = u_{yy}$ is subjected to a process of reciprocation.[2]

According to Fourier's inversion formula, the density of probability assigned, for a fixed $\lambda > 0$, by the (even) Fourier transform $e^{-|t|^\lambda}$, $-\infty < t < \infty$, is $\pi^{-1}f_\lambda(x)$, $-\infty < x < \infty$, where

$$f_\lambda(x) = \int_0^\infty e^{-t^\lambda} \cos(xt)\, dt \equiv f_\lambda(-x); \qquad 0 \leqq x < \infty. \tag{1}$$

Partial integration of (1) gives

$$f_\lambda(x) = \lambda \int_0^\infty e^{-t^\lambda} t^\lambda S(xt)\, dt, \tag{2}$$

where $S(u) = \sin u/u$. Hence, if t is replaced by $(t/x)^{1/\lambda}$ for a fixed $x > 0$,

$$xf_\lambda(x) = x^{-\lambda} g_\lambda(x^{-\lambda}), \qquad (x > 0, x^{-\lambda} > 0), \tag{3}$$

where

$$g_\lambda(z) = \int_0^\infty e^{-zt} \sin t^{1/\lambda}\, dt. \tag{4}$$

All derivatives $f_\lambda^{(n)}(x)$ exist for $-\infty < x < \infty$, since the integrals

$$(-1)^n \int_0^\infty e^{-t^\lambda} t^{2n+1} \sin(xt)\, dt, \qquad (-1)^n \int_0^\infty e^{-t^\lambda} t^{2n} \cos(xt)\, dt,$$

Received July 11, 1941. The author is a Fellow of the Guggenheim Foundation.

[1] A. Cauchy, *Oeuvres Complètes*, ser. 1, vol. 12(1900), pp. 94–114. Today it is known that, the density of probability being required to be non-negative, Cauchy's distributions exist if and only if λ does not exceed the value $\lambda = 2$ belonging to the limiting case of a normal distribution; cf. P. Lévy, *Calcul des probabilités*, Paris, 1925, pp. 252–277. For sharper results, see A. Wintner, *On a class of Fourier transforms*, American Journal of Mathematics, vol. 58(1936), pp. 45–90.

[2] Cf. P. Lévy, *Sur certains processus stochastiques homogènes*, Compositio Mathematica, vol. 7(1939), pp. 283–339.

obtained by formal differentiation of the integral (1), are uniformly convergent for $-\infty < x < \infty$. In particular

$$f_\lambda^{(2n+1)}(0) = 0, \qquad f_\lambda^{(2n)}(0) = (-1)^n \int_0^\infty e^{-t^\lambda} t^{2n}\, dt.$$

Since, if t^λ is replaced by t, the last integral becomes $\lambda^{-1}\Gamma([2n+1]/\lambda)$, the Taylor series of f_λ at the origin is

$$(5) \qquad \sum_{n=0}^{\infty} f_\lambda^{(n)}(0) z^n/n! = \lambda^{-1} \sum_{n=0}^{\infty} (-1)^n \Gamma([2n+1]/\lambda) z^{2n}/\Gamma(2n+1).$$

It is clear from Stirling's formula that this power series diverges for every $z \neq 0$ unless $\lambda \geqq 1$.

The results to be proved imply the following theorem:

If the index of a symmetric stable distribution function is less than 1, *the density of probability defines an analytic function possessing a branch point at the origin; a branch point which represents a transcendental singularity and is the only finite singularity of the analytic function.*

Replace x by $z = x + iy$ in (1); so that

$$(6) \qquad f_\lambda(z) = \int_0^\infty e^{-t^\lambda} \cos(zt)\, dt.$$

Since $|\cos(zt)|$, where $t \geqq 0$, is not greater than $e^{|z|t}$, the integral (6) is uniformly convergent on every fixed z-circle in the z-plane, if $\lambda > 1$. If $\lambda = 1$, the integral (6) represents $(1 + z^2)^{-1}$ in the strip $-1 < y < 1$. Thus $f_\lambda(z)$ is a transcendental entire function or a rational function (with simple poles at $z = \pm i$), according as $\lambda > 1$ or $\lambda = 1$.

It will from now on be assumed that $0 < \lambda < 1$. Then the remark following (5) implies that $f_\lambda(z)$ cannot be regular at $z = 0$. It also is clear from $0 < \lambda < 1$ that the integral (6) is divergent whenever $z \neq x$, where $z = x + iy$. However, (3) supplies an analytic continuation of (1) for certain complex z.

In fact, if $z = x + iy$ is arbitrary, an application of the second mean-value theorem to the real and imaginary parts shows that, if $x \geqq 0$, $b > a$,

$$\left| \int_a^b e^{-zt} \sin t^{1/\lambda}\, dt \right| \leqq 4e^{-za} \max_{a \leqq c \leqq b} \left\{ \left| \int_a^c \sin t^{1/\lambda}\, dt \right| + \left| \int_c^b \sin t^{1/\lambda}\, dt \right| \right\}.$$

Hence, for every $b > a$ and uniformly for $x \geqq 0$, $-\infty < y < \infty$,

$$\left| \int_a^b e^{-zt} \sin t^{1/\lambda}\, dt \right| \leqq 12 \max_{a \leqq x < \infty} \left| \int_a^x \sin t^{1/\lambda}\, dt \right| \to 0 \quad \text{as} \quad a \to \infty,$$

since $0 < \lambda < 1$. Accordingly, the integral (4) is uniformly convergent in the closed half-plane $x \geqq 0$. Hence, the function $g_\lambda(z)$ is continuous in this half-plane and is regular analytic in the open half-plane $x > 0$.

On the other hand, it is clear that the Laplace transform (4) is divergent for every negative z, and therefore in the whole half-plane $x < 0$. Nevertheless,

the function $g_\lambda(z)$ represented by this integral in the half-plane $x \geqq 0$ fails to have a singularity on the line $x = 0$. In fact, $g_\lambda(z)$ can be developed for *every* z into the power series

$$(7) \qquad \sum_{n=0}^{\infty} g_\lambda^{(n)}(0)z^n/n! = \lambda \sum_{n=0}^{\infty} (-1)^n\Gamma([n+1]\lambda) \sin(\tfrac{1}{2}\pi[n+1]\lambda)z^n/\Gamma(n+1).$$

To this end, let α, r and $\lambda < 1$ be positive constants, and let

$$\int_C F(z)\,dz = 0, \qquad F(z) = z^{\alpha-1}e^{iz-rz^\lambda}, \qquad (z^\mu > 0 \text{ for } z > 0),$$

be applied to the contour C consisting of the two segments and the two arcs

$$z = t, \qquad z = it, \qquad z = \epsilon e^{i\phi}, \qquad z = Re^{i\phi}, \qquad (0 < \epsilon < R < \infty),$$

where $\epsilon \leqq t \leqq R$ and $0 \leqq \phi \leqq \frac{1}{2}\pi$. Since $\alpha > 0$, $r > 0$ and $0 < \lambda < 1$, it is readily ascertained that the contribution of the two arcs tends to 0 as $\epsilon \to 0$ and $R \to \infty$. Hence, the integrals of $F(t)$ and $iF(it)$ over the half-line $0 < t < \infty$ have a common value. Thus, by the definition of F,

$$(8) \qquad \int_0^\infty e^{-rt^\lambda}t^{\alpha-1}e^{it}\,dt = e^{\frac{1}{2}\pi\alpha i}\int_0^\infty e^{-t-rt^\lambda\cos\frac{1}{2}\pi\lambda}t^{\alpha-1}e^{-irt^\lambda\sin\frac{1}{2}\pi\lambda}\,dt.$$

On the other hand, since $\lambda > 0$, the n-th derivative of the analytic function represented by (4) in the half-plane $x > 0$ may be obtained by differentiating (4) beneath the integral sign, if $x > 0$, where $z = x + iy$. Thus, for $z = r > 0$,

$$g_\lambda^{(n)}(r) = (-1)^n\int_0^\infty e^{-rt}t^n\sin t^{1/\lambda}\,dt \equiv (-1)^n\lambda\int_0^\infty e^{-rt^\lambda}t^{(n+1)\lambda-1}\sin t\,dt,$$

if t is replaced by t^λ. Hence, by the imaginary part of (8),

$$g_\lambda^{(n)}(r) = (-1)^n\lambda\int_0^\infty e^{-t-rt^\lambda\cos\frac{1}{2}\pi\lambda}t^{(n+1)\lambda-1}\sin(\tfrac{1}{2}\pi[n+1]\lambda - rt^\lambda\sin\tfrac{1}{2}\pi\lambda)\,dt,$$

where $(n+1)\lambda = \alpha$ and $0 < r < \infty$. Since $0 < \lambda < 1$, this implies that

$$(9) \qquad |g_\lambda^{(n)}(r)| < \lambda\int_0^\infty e^{-t}t^{(n+1)\lambda-1}\,dt \equiv \lambda\Gamma([n+1]\lambda) \qquad \text{for} \qquad 0 < r < \infty,$$

and that $g_\lambda^{(n)}(r)$ tends, as $r \to +0$, to the limit

$$(-1)^n\lambda\int_0^\infty e^{-t}t^{(n+1)\lambda-1}\sin(\tfrac{1}{2}\pi[n+1]\lambda)\,dt \equiv (-1)^n\lambda\Gamma([n+1]\lambda)\sin(\tfrac{1}{2}\pi[n+1]\lambda)$$

for every n (so that, by Rolle's theorem, the derivative $g_\lambda^{(n)}(0)$, when thought of as a right-hand derivative, exists and is represented by

$$(10) \qquad g_\lambda^{(n)}(0) = (-1)^n\lambda\Gamma([n+1]\lambda)\sin(\tfrac{1}{2}\pi[n+1]\lambda)$$

for every n).

Since the function (4) is regular analytic in the half-plane $x > 0$, it can be developed in the neighborhood of every $z = r > 0$ into a power series,

$$g_\lambda(z) = \sum_{n=0}^{\infty} g_\lambda^{(n)}(r)(z - r)^n/\Gamma(n + 1),$$

having a radius of convergence not less than r. But (9) shows that this power series is dominated by one which, in view of Stirling's formula, is convergent in the whole $(z - r)$-plane, since $0 < \lambda < 1$. Consequently, the radius of convergence of the expansion of $g_\lambda(z)$ according to the powers of $z - r$ is infinite (for every $r > 0$). This proves that $g_\lambda(z)$ is an entire function. Hence, the expansion (7) of $g_\lambda(z)$ follows from (10) for every z.

Since $g_\lambda(z)$ is an entire function, it follows from (3), where $x^{-\lambda} = 1/x^\lambda$, that the analytic continuation $f_\lambda(z)$ of the real function $f_\lambda(x)$ represented on the half-line $\arg z = 0$ by the integral (1) is a regular function of the position on the Riemann surface of $\log z$ or of a real rational power of z according as λ is irrational or rational, where $0 < \lambda < 1$. However, the branch point $z = 0$ does not belong to the Riemann surface of $f_\lambda(z)$ even if λ is rational, since the singularity of f_λ at $z = 0$ is transcendental. This is implied by the remark following (5), and also by (3), since the entire function $g_\lambda(z)$ is not a polynomial. It also is seen from (3) that, notwithstanding the even character of the real Fourier transform (1), the analytic continuations of the two real functions $f_\lambda(x)$, $0 < x < \infty$ and $f_\lambda(x)$, $-\infty < x < 0$ could belong to the same analytic function only by an accident. This case can easily be discussed.

Since $\Gamma(s)\Gamma(1 - s) = \pi/\sin \pi s$, there is a striking reciprocity between the expansions (5), (7). Correspondingly, while the functions (4), (6) are connected by the relation (3) for every $\lambda > 0$, the power series (5), (7) converge for a common value of λ and for some $z \neq 0$ only in the trivial case $\lambda = 1$ (in which case both series reduce to $1 - z^2 + z^4 - \cdots = (1 + z^2)^{-1}$, if $|z| < 1$). In fact, (5), (7) diverge for every $z \neq 0$ whenever $\lambda < 1$, $\lambda > 1$ respectively; while (5), (7) obviously are entire functions of the respective orders $(1 - \lambda^{-1})^{-1}$, $(1 - \lambda)^{-1}$ in the complementary cases $\lambda > 1$, $\lambda < 1$.

It should finally be mentioned that (7) is formally related to the standard entire functions occurring in the theories of explicit analytic continuation beyond the circle of convergence of a power series.

THE JOHNS HOPKINS UNIVERSITY.

CONTINUA OF FINITE SECTIONS

By O. G. Harrold, Jr.

Introduction. The problem has been proposed to characterize the continua on which a real valued continuous function of finite sections can be defined, i.e., a real valued continuous function f such that for each y, $f^{-1}(y)$ is a finite set of points [1].[1] Čech gave three necessary conditions which such a continuum satisfies. First, the continuum M is regular in the Menger-Urysohn sense, second, M has at most a countable number of end-points and, third, M has no continuum of condensation. Mazurkiewicz considered and solved the same problem for the class of dendrites [2]. A dendrite has the above stated property $\mathfrak{C}$ if and only if the set of end-points is countable and the operation of taking the bi-lateral coherence of the set of ramification points produces the null set on iteration.[2] Aitchison [3], Eilenberg [4], G. T. Whyburn [5], and the author [6] have also used the method of mappings into an interval to characterize types of continua.

In this paper we give a solution of the problem proposed above.

Theorem. *The continuum M has the property $\mathfrak{C}$ if and only if M is locally connected and every dendrite in M has the property.*

The necessity of the conditions is obvious in view of the previously stated results and the remark that if M has the property every subcontinuum also has it.

The sufficiency will follow from the assertions (1)–(5) below.

Let M denote a Peano continuum such that every dendrite in M has the property $\mathfrak{C}$.

(1). *M has no continuum of condensation.*

(2). *There exists a dendrite D in M which contains the enclosure of the set of points $X \subset M$, where $x \in X$ if and only if the Menger order of $x \neq 2$.*

(3). *There exists a map $f(M) = (0, 1)$ such that $f^{-1}(y)$ is finite for all but a countable set of values y and for these exceptional values $f^{-1}(y)$ is countable.*

Received July 13, 1941. This work was started while the author was a National Research Fellow at the University of Virginia.

[1] The numbers in brackets refer to the bibliography at the end of the paper.

[2] The bi-lateral coherence $\Phi(U)$ of a subset U of a dendrite D is defined as follows. $\Phi(U)$ is the set of all points $x \in D$ such that for at least two components T_1, T_2 of $D - x$, $x \in (T_1 \cdot U)' \cdot (T_2 \cdot U)'$. Set $\Phi_0(U) = U$, and

$$\Phi_\alpha(U) = \Phi[\Phi_{\alpha-1}(U)], \prod_{i<\alpha} \Phi_i(U)$$

according as the ordinal α has a predecessor or not. In case U is countable it is clear that there is a first ordinal α of the first or second number class such that $\Phi_{\alpha+1}(U) = \Phi_\alpha(U)$. See [2].

(4). *If Δ is the set of exceptional values in* (3), *every subset of Δ has isolated points.*

(5). *There exists a map $g(M) = (0, 1)$ which is obtained by modification of f in the neighborhood of $f^{-1}(y)$, $y \in \Delta$, such that $g^{-1}(y)$ is finite for all y.*

Proof of (1). Suppose T is a continuum such that $T \subset \overline{M - T}$. Since every dendrite has property C and therefore has at most a countable number of end-points, M is a hereditary arc sum [7]. Hence T may be taken to be an arc. Since T is a continuum of condensation of M, there exists a sequence of arcs (Y_i) such that $Y_i \cdot Y_j = 0$, $i \neq j$, $Y_i \cdot T = y_i$, a point, and $T = \overline{\sum y_i}$. The set $Z = T + \sum Y_i$ is a dendrite in M with a continuum of condensation and therefore cannot have property C.

Proof of (2). If M has no continuum of condensation, $\bar{X}$ is totally disconnected [8]. Since M is Peanian, there is a dendrite D such that $\bar{X} \subset D \subset M$. Each component U_i of $M - D$ is an open free arc whose enclosure is either an arc or a simple closed curve. There is no loss in assuming that the dendrite D has been so selected that each component U_i of $M - D$ has as enclosure an arc $\bar{U}_i$ which lies wholly in the interior of a free arc of M.

Proof of (3). By our hypothesis, there exists a map $f(D) = (0, 1)$ which is of finite sections, i.e., such that each $f^{-1}(y)$ is finite. Let the end-points of U_i be a_i and b_i. It may be supposed that $f(a_i)$ is a rational number while $f(b_i)$ is irrational, for f is constant on no subcontinuum of D and we may truncate a portion of the free arc of D containing a_i (say) to obtain a dendrite D^0 satisfying (2) and such that the enclosure of every component of $M - D^0$ is an arc lying in a free arc in M. Let D^1 denote the dendrite which is obtained by performing this alteration for all the sets U_i. Thus f is a map of finite sections on the dendrite $D^1 \supset \bar{X}$. The enclosure of each component U_i of $M - D^1$ is an arc lying wholly in a free arc of M and if a_i and b_i are the end-points of $\bar{U}_i$, $f(a_i) \neq f(b_i)$.

The map f will now be extended to M by what we shall refer to as a *normal extension*. The extended mapping will satisfy (3). By the choice of D^1, $f(a_i) \neq f(b_i)$. Hence on $\bar{U}_i$ the mapping f can be extended so as to map $\bar{U}_i$ topologically into the interval determined by $f(a_i)$ and $f(b_i)$. For any y, $f^{-1}(y) = f^{-1}(y) \cdot D^1 + \sum_1^{\infty} f^{-1}(y) \cdot U_i$. Thus $f^{-1}(y)$ is countable. Let Δ be the set of values y for which $f^{-1}(y)$ is infinite. To complete the proof of (3) it must be shown that Δ is countable. Let $z \in [f^{-1}(y)]'$. Clearly, $z \in D^1$. It will be shown first that z is necessarily an im kleinen cycle point of M, i.e., a point situated on simple closed curves in M of arbitrarily small diameter. There exists a subsequence (U_{k_i}) of (U_i) such that $U_{k_i} \cdot f^{-1}(y) = z_i$, $z_i \to z$. All points of $f^{-1}(y)$ sufficiently near z may be supposed to be in the sequence (z_i), since z is an isolated point of $[f^{-1}(y)]'$. Let F be a region in D^1 containing z such that $f^{-1}(y) \cdot F = z$. For $i \geqq i_0$, a_{k_i} and b_{k_i} lie in F. Set $x_i = f(a_{k_i})$, $y_i = f(b_{k_i})$. If $y = f(z)$ separates only a finite number of the pairs of points x_i, y_i on the interval $(0, 1)$, then $z \notin [f^{-1}(y)]'$. Hence there is a subsequence

x_{j_i}, y_{j_i} such that for each i either $x_{j_i} < y < y_{j_i}$ or $y_{j_i} < y < x_{j_i}$. Since z is the only inverse to y in F, z must cut F between a_{j_i} and b_{j_i}. Hence there exist arcs (in F) X_{j_i}, Y_{j_i} from z to x_{j_i}, y_{j_i}, respectively, such that $X_{j_i} \cdot Y_{j_i} = z$. The set $X_{j_i} + Y_{j_i} + U_{j_i}$ is a simple closed curve in M containing z. Thus z is an im kleinen cycle point.

Since M is a hereditary arc sum, M has only a countable number of non-local separating points [7]. Thus if Δ were uncountable, there would be uncountably many values y each of whose inverse sets would consist entirely of local separating points. But we have just shown that each such inverse set contains an im kleinen cycle point. But every im kleinen cycle point which is also a local separating point is a ramification point, i.e., of order greater than 2 [9]. By the local separating point order theorem only a countable number of such points can exist in any continuum [10]. Hence Δ is countable and (3) is established.

Proof of (4). Suppose, on the contrary, that Δ contains a dense-in-itself subset Δ_1. For $0 \leqq y \leqq 1$, let $m(y)$ be the number of $f^{-1}(y)$ in $\sum U_i$. For every positive integer n the set of all y such that $m(y) > n$ is an open set B_n on (0, 1). The common part of $B_1, B_2, B_3, \cdots$ is Δ. Therefore Δ is a G_δ set containing a dense-in-itself subset Δ_1. It follows by Baire's theorem that Δ contains a perfect set, and is therefore uncountable. Thus (4) is established.

Since the number of im kleinen cycle points in M is countable, a homeomorphism h of (0, 1) into itself may be defined such that for each im kleinen cycle point x, $g(x) = h[f(x)]$ is a rational number. If for each $t \in M$ we set $g(t) = h[f(t)]$, a mapping of finite sections on D^1 is obtained. Let now D be any dendrite in M such that $\bar{X} \subset D \subset D^1$ and for the end-points a_i, b_i of a component U_i of $M - D$, $g(a_i)$ is rational and $g(b_i)$ is irrational. The normal extension of g from D to M exists and the corresponding Δ satisfies (3) and (4). Further, $y \in \Delta$ implies y is rational.

We shall have occasion to define certain piecewise homeomorphisms of (0, 1) into (0, 1). These mappings will always be thought of as linear with rational coefficients. Thus the properties of $g(a_i)$ rational, $g(b_i)$ irrational, and $y \in \Delta$ implying y rational will be preserved. The symbol f will be retained for a mapping from which a modified mapping is to be formed.

Before proceeding with the proof of (5), it will be helpful to have the following sequence of lemmas.

LEMMA A. *Let Δ consist of the point $y = \frac{1}{2}$. A modified map g can be defined on D such that for the normal extension of g, Δ is vacuous. Further,* $f^{-1}(i) = g^{-1}(i)$, $i = 0, 1$.

Case 1. The set $D \cdot f^{-1}(1/2)$ consists of the single point x^1. Let T_1 denote a linear map of the interval (0, 1/3) on (0, 1/2) with the point 0 fixed. Let T_2 denote a linear map of (2/3, 1) on (1/2, 1) with 1 fixed. Let T denote the map obtained by considering T_1 and T_2 acting simultaneously. If $y \leqq 1/3$ or $\geqq 2/3$, set $g(x) = T[f(x)]$. Let Q denote $1/3 < y < 2/3$. If $x \in W$, where W is a component of $f^{-1}(Q)$, and $f(\bar{W} - W) = 1/3$, let $Z(x)$ denote the symmetrically

located point to $f(x)$ with respect to $y = 1/3$. If $f(\bar{W} - W) = 2/3$, let $Z(x)$ denote the point symmetric to $f(x)$ with respect to $2/3$. For $x \in W$, define $g(x) = T[Z(x)]$. It now remains to define $g(x)$ on the component C of $f^{-1}(Q)$ containing x^1.

Since $x^1 \in D \cdot f^{-1}(\Delta)$, there are infinitely many components U_i of $M - D$ such that $y = 1/2$ separates $f(a_i)$ and $f(b_i)$. Let A be the subset of $\bar{C}$ which is obtained by deleting all open arcs U_i such that $U_i \subset C$. If $1/3 \leqq y \leqq 1/2$, set $y = 1/2 - \epsilon$, and define $H(y) = 1/2 + \epsilon$. If $1/2 \leqq y \leqq 2/3$, $H(y) = y$. Set $g(x) = H[f(x)] - 1/6$, $x \in A$. By the choice of D, H and f, $g(a_i) \neq g(b_i)$, thus a normal extension of g may be defined on $\bar{C}$ such that $g(\bar{C}) = (1/3, 1/2)$. For each arc U_i added to A to form $\bar{C}$ we have $g(a_i), g(b_i) > 1/3$. Hence the extension of g to $\bar{C}$ does not add any points which map into $1/3$ and only a finite number of points inverse to any $y > 1/3$. Thus for all y, $g^{-1}(y) \cdot \bar{C}$ is finite. Now g is defined over all of M. If $y = g(x) < 1/3$ or $> 1/2$, $g^{-1}(y) \subset f^{-1}T^{-1}(y) + f^{-1}Z^{-1}T^{-1}(y)$. If $1/3 \leqq y \leqq 1/2$, $g^{-1}(y) \subset f^{-1}T^{-1}(y) + f^{-1}Z^{-1}T^{-1}(y) + f^{-1}H^{-1}(y + 1/6) + g^{-1}(y) \cdot \bar{C}$. Thus $g^{-1}(y)$ is finite in any case. Clearly, $g^{-1}(i) = f^{-1}(i)$, $i = 0, 1$.

Case 2. $D \cdot f^{-1}(\Delta) = x_1 + x_2 + \cdots + x_n$. Since f is constant on no subcontinuum of M, given any $d > 0$ there exists an $e > 0$ such that the diameter of any component of the inverse of an interval of $(0, 1)$ of diameter $< e$ will be $< d$. Let $d = 1/2 \min \rho(x_i, x_j)$, $i \neq j$. Let J be an interval of $(0, 1)$ containing $1/2$ in its interior and of diameter $< e < 1/2$. Only a finite number of the components of $f^{-1}(J)$ can intersect $f^{-1}(1/2)$ and no component contains two x_i's. Let the components of $f^{-1}(J)$ intersecting $f^{-1}(1/2)$ be $C^1, C^2, \cdots, C^k$. Suppose $C^1, C^2, \cdots, C^m$, $m \leqq k$, contain points of $[f^{-1}(1/2)]'$. For $i \leqq m$, $f(C^i)$ contains $1/2$ in its interior and the results of Case 1 can be used to redefine f on $\bar{C}^i$ so that f has finite sections on $\bar{C}^i$ and f is unaltered on $\bar{C}^i - C^i$. The application of this principle to each $\bar{C}^i$, $i \leqq m$, with f fixed on $M - (C^1 + C^2 + \cdots + C^m)$ yields a map $g(M) = (0, 1)$ which is of finite sections and $g^{-1}(i) = f^{-1}(i)$, $i = 0, 1$. This completes the proof of Lemma A.

LEMMA B. *If each point of* Δ *is isolated* ($\Delta \cdot \Delta' = 0$), *there exists a map* $g(M) = (0, 1)$ *of finite sections with* $g^{-1}(i) = f^{-1}(i)$, $i = 0, 1$.

Since each point of Δ is isolated and neither 0 nor 1 is in Δ, there is a sequence of subintervals (I_i) of $(0, 1)$ such that $I_i \cdot I_j = 0$, $i \neq j$, y_i is in the interior of I_i, where $\Delta = \sum_1^{\infty} y_i$, and each I_i is in the interior of $I = (0, 1)$. Let $g_1(x)$ be a function obtained by applying Lemma A to each of the components V of $f^{-1}(I_i)$ which contain points of $[f^{-1}(y_1)]'$, where the numbers 0 and 1 of the Lemma are replaced by the numbers corresponding to the end-points of $f(V)$. On the complement of the sum of these components in M set $g_1(x) = f(x)$. Repeating this procedure, a sequence of functions $g_i(x)$ is defined with $g_i(x) = g_{i+1}(x)$ except on the components of $f^{-1}(I_{i+1})$ which contain points of $[f^{-1}(y_{i+1})]'$. For $y \in I - (y_{i+1} + y_{i+2} + \cdots)$, $g_i^{-1}(y)$ is finite. For each $x \in M$, set $g(x) =$

$\lim g_i(x)$. Since $g_i(x)$, $x \in f^{-1}(I_j)$, is fixed for $i \geqq j$, and $g = f$ on $M - \sum_1^\infty f^{-1}(W_i)$, where W_i is the interior of I_i, g is continuous. It is readily verified that g is of finite sections. Since 0 and 1 are in no I_i, $f^{-1}(i) = g^{-1}(i)$, $i = 0, 1$.

LEMMA C. *If $(\Delta \cdot \Delta')' = 0$, a map g of finite sections exists with $g^{-1}(1) = f^{-1}(1)$.*

Let $\Delta \cdot \Delta'$ consist of the points y_i, $i = 1, 2, \cdots, n$, where $0 < y_1 < y_2 < \cdots < y_n < 1$. Let z_i, $i = 1, 2, \cdots, n$, be n rational numbers such that $0 < z_1 < y_1 < z_2 < y_2 < \cdots < z_n < y_n < 1$. Let T denote the continuous transformation resulting by mapping in piecewise linear fashion $(0z_1)$ on $(0z_1)$ with z_1 fixed, (z_1y_1) on $(0z_1)$ with z_1 fixed, (y_1z_2) on $(0z_2)$ with z_2 fixed, etc., finally, (y_n1) on $(0, 1)$ with 1 fixed. The existence of such piecewise linear maps utilizes the assumption that $y \in \Delta$ implies y rational. On the dendrite D set $h(x) = T[f(x)]$. Clearly, $h(a_i) \neq h(b_i)$. Hence the normal extension of h to M exists. (The map h is of finite sections on D.) Denote the extended map by g.

It will next be shown that $\Delta_g \cdot \Delta_g' = 0$, so that Lemma B will be applicable. It is noted first that the points of $f^{-1}(y_i)$, $i = 1, 2, \cdots, n$, are points at which $h = g$ has the minimum value 0 and hence on account of the nature of a normal extension, $0 \,\epsilon'\, \Delta_g$. Next, $y \in \Delta_g$ is equivalent to $T^{-1}(y) \cdot [\Delta - (y_1 + y_2 + \cdots + y_n)] \neq 0$. If $t \in T^{-1}(y) \cdot [\Delta - (y_1 + \cdots + y_n)]$, $f^{-1}(t)$ is infinite; hence $g^{-1}[T(t)] \supset f^{-1}(t)$ implies $y \in \Delta_g$. If $y \in \Delta_g$, some point t of $T^{-1}(y)$ must have $f^{-1}(t)$ infinite since $T^{-1}(y)$ is finite. Since $T(y_1 + \cdots + y_n) = 0$, $T^{-1}(y) \cdot [\Delta - (y_1 + \cdots + y_n)] \neq 0$. But about each of the points of the finite set $T^{-1}(y) \cdot [\Delta - (y_1 + \cdots + y_n)]$ there is a neighborhood V such that $V \cdot [\Delta - T^{-1}(y)] = 0$. Clearly, the common part of the transforms (by T) of these neighborhoods contains a neighborhood of y such that it contains no other point of Δ_g. Applying the Lemma B to the function g we get the stated result. It is noted that for the new g we can claim only $g^{-1}(1) = f^{-1}(1)$; however this is sufficient for our purposes.

Put $\Delta_1 = \Delta \cdot \Delta'$. If the ordinal α has a predecessor, set $\Delta_\alpha = \Delta_{\alpha-1} \cdot \Delta_{\alpha-1}'$. If α is a limit ordinal, set $\Delta_\alpha = \prod_{i<\alpha} \Delta_i$.

LEMMA D. *If for all ordinals $i < \alpha$ the existence of a map $f(M) = (0, 1)$ with $\Delta_i = 0$ implies that a function g of finite sections exists on M with $g^{-1}(1) = f^{-1}(1)$, then $f(M) = (0, 1)$ with $\Delta_\alpha = 0$ implies the existence of such a map.*

Case 1. The ordinal α has a predecessor, so that $\Delta_\alpha = \Delta_{\alpha-1} \cdot \Delta_{\alpha-1}' = 0$. Set $\Delta_{\alpha-1} = \sum_1^\infty y_i$. Let (I_i) be a sequence of pairwise disjoint intervals such that y_i lies in the interior of I_i and the end-points of I_i lie in $I - (\Delta + 0 + 1)$. Let W_i be the interior of I_i. Let C_λ^i, $\lambda = 1, 2, \cdots, n$, be the components of $f^{-1}(W_i)$ intersecting $f^{-1}(y_i)$. Let V_λ^i, $\lambda = 1, 2, \cdots$, be the remaining components. On the continua $\bar{V}_\lambda^i$ the inductive assertion holds. Hence there is a map $g_{i\lambda}(V_\lambda^i) = f(V_\lambda^i)$, $\lambda = 1, 2, \cdots$, of finite sections and if $y \in f(\bar{V}^i) \cdot (I_i - W_i)$, $(g_{i\lambda})^{-1}(y) = f^{-1}(y) \cdot \bar{V}_\lambda^i$. The point y_i is the only point of $\Delta_{\alpha-1}$ in I_i. If we replace 0, 1 and $\frac{1}{2}$ by the numbers corresponding to the end-points of I_i and y_i

respectively, the procedure of Lemma A gives a map $f_{i\lambda}(\bar{C}^i_\lambda) = f(\bar{C}^i_\lambda)$ such that $\Delta_{\alpha-1} = 0$ for $f_{i\lambda}$. Hence the inductive assumption applies to $\bar{C}^i_\lambda$ and there exists a map $h_{i\lambda}(\bar{C}^i_\lambda) = f(\bar{C}^i_\lambda)$ which is of finite sections and for $x \in \bar{C}^i - C^i$, $h_{i\lambda}(x) = f(x)$. On $f^{-1}(I_i)$ set $g_i(x) = g_{i\lambda}(x)$, $x \in V^i_\lambda$. If $x \in \bar{C}^i_\lambda$, set $g_i(x) = h_{i\lambda}(x)$. We observe that since the sets $\bar{C}^i_\lambda$ are finite in number and the sets $\bar{V}^i_\lambda$ form a null sequence, for any $y \in W_i$, $g_i^{-1}(y)$ is finite. We prove next that if $y \in I_i - W_i$, then $g_i^{-1}(y)$ is finite. Clearly, $g_i^{-1}(y) \subset f^{-1}(y) + h_{i1}^{-1}(y) + \cdots + h_{in}^{-1}(y) + \sum_{\lambda=1}^{\infty} g_{i\lambda}^{-1}(y)$. But one of the properties of the $g_{i\lambda}$ is that $g_{i\lambda}^{-1}(y) \subset f^{-1}(y)\cdot\bar{V}^i$. Hence the collection of points on the right is finite. Thus $g_i^{-1}(y)$ is finite. On $M - \sum f^{-1}(W_i)$ set $g(x) = f(x)$. On $f^{-1}(W_i)$ set $g(x) = g_i(x)$. The function g is continuous since it is continuous on $f^{-1}(W_i)$ and agrees with f on $\sum f^{-1}(I_i - W_i)$. The map g is of finite sections.

Case 2. α is a limit ordinal, thus $\Delta_\alpha = \prod_{i<\alpha} \Delta_i = 0$. Set $\Delta = \sum_1^{\infty} y_i$. To y_1 there is a first ordinal $\beta_1 < \alpha$ such that $y_1 \in' \Delta_{\beta_1+1}$. Hence y_1 is an isolated point of Δ_{β_1}. Let I_1 be an interval containing y_1 in its interior with $I_1\cdot\Delta_{\beta_1} = y_1$ and the end-points of I_1 are in $I - (\Delta + 0 + 1)$. Let y_{k_2} be the first point of Δ in $I - I_1$. To y_{k_2} there is a first ordinal $\beta_{k_2} < \alpha$ such that $y_{k_2} \in' \Delta_{\beta_{k_2}+1}$. Thus y_{k_2} is an isolated point of $\Delta_{\beta_{k_2}}$. Let I_2 be an interval containing y_{k_2} in its interior such that (a) $I_1\cdot I_2 = 0$, (b) $I_2\cdot\Delta_{\beta_{k_2}} = y_{k_2}$ and (c) I_2 has its end-points in $I - (\Delta + 0 + 1)$. Continuing in this way there is determined a sequence of intervals (I_i) such that (a) $I_i\cdot I_j = 0$, $i \neq j$, (b) $I_i\cdot\Delta_{\beta_{k_i}} = y_{k_i}$, (c) I_i has end-points in $I - (\Delta + 0 + 1)$, and (d) by the choice of the y_{k_i}, $\Delta \subset \sum_1^{\infty} I_i$.

Let W_i denote the interior of I_i and V^i_j, $j = 1, 2, \cdots, n$, those components of $f^{-1}(W_i)$ such that $f(\bar{V}^i_j) = I_i$. Let the other components be Z^i_j, $j = 1, 2, \cdots$. Our inductive assertions hold for each $\bar{Z}^i_j$ and $\bar{V}^i_j$ since $\beta_{k_i} < \alpha$. Let z_{ij} be a map of finite sections on $\bar{Z}^i_j$ such that $z_{ij}(\bar{Z}^i_j) = f(\bar{Z}^i_j)$ and if $y \in (I_i - W_i)\cdot z_{ij}(Z^i_j)$, then $z_{ij}^{-1}(y) = f^{-1}(y)\cdot\bar{Z}^i_j$. Let v_{ij} be a map of finite sections on $\bar{V}^i_j$ such that $v_{ij}(\bar{V}^i_j) = I_i$. Here we suppose that f and v_{ij} agree on $f^{-1}(I_i - W_i)$, but we cannot (and do not need to) assert that if $y \in (I_i - W_i)v_{ij}(\bar{V}^i_j)$, then $v_{ij}^{-1}(y) = f^{-1}(y)\bar{V}^i_j$.

On $x \in V^i_j$ set $g(x) = v_{ij}(x)$. On $x \in Z^i_j$ set $g(x) = z_{ij}(x)$ and on $x \in M - \sum_1^{\infty} f^{-1}(W_i)$ set $g(x) = f(x)$. The continuity of g is clear, as well as $g^{-1}(1) = f^{-1}(1)$. It remains only to verify that $g^{-1}(y)$ is finite for $y \in I_i$. We have $g^{-1}(y) \subset f^{-1}(y) + \sum_{j=1}^{n} v_{ij}^{-1}(y) + \sum_{j=1}^{\infty} z_{ij}^{-1}(y)$. If $y \in W_i$, only a finite number of the sets $z_{ij}^{-1}(y)$ can be non-vacuous since the sets Z^i_j form a null family. Hence $g^{-1}(y)$ is finite. If $y \in I_i - W_i$, the condition $z_{ij}^{-1}(y) \subset f^{-1}(y)$ is fulfilled. Hence again $g^{-1}(y)$ is finite. Hence $g(M) = (0, 1)$ has finite sections. This completes the proof of Lemma D.

Proof of (5). For all ordinals α of the first or second number class $\Delta_{\alpha+1} \subset \Delta_\alpha$. Since Δ is countable, there is an ordinal α_0 such that $\Delta_{\alpha_0} = \Delta_{\alpha_0+1}$. By virtue of (4), $\Delta_{\alpha_0} = 0$. Hence application of Lemma D yields the result that there is a map $g(M) = (0, 1)$ of finite sections.

NORTHWESTERN UNIVERSITY.

BIBLIOGRAPHY

1. E. ČECH, *Une nouvelle classe de continua*, Fundamenta Mathematicae, vol. 18(1932), pp. 86–88.
2. S. MAZURKIEWICZ, *Une classe de dendrites*, Fundamenta Mathematicae, vol. 18(1932), pp. 89–98.
3. B. AITCHISON, *On the mapping of locally connected continua into arcs*, Comptes Rendus des Séances de la Société des Sciences et des Lettres de Varsovie, vol. 27(1934), III, pp. 3–19.
4. S. EILENBERG, *On continua of finite length*, Annales de la Société Polonaise de Mathematiques, vol. 17(1938), pp. 253–254.
5. G. T. WHYBURN, *Characterizations of certain curves by continuous functions defined upon them*, American Journal of Mathematics, vol. 55(1933), pp. 131–134.
6. O. G. HARROLD, *The rôle of local separating points in certain problems of continuum structure*, Proceedings of the University of Michigan Conference on Topology (1941).
7. O. G. HARROLD, *Hereditary arc sums*, Duke Mathematical Journal, vol. 5(1939), pp. 111–117.
8. K. MENGER, *Kurventheorie*, Teubner, Berlin (1932).
9. G. T. WHYBURN, *Concerning points of continuous curves defined by certain im kleinen properties*, Mathematische Annalen, vol. 102(1929), pp. 313–336.
10. G. T. WHYBURN, *Local separating points of continua*, Monatsheften für Mathematik und Physik, vol. 36(1929), pp. 305–414.

THE COEFFICIENTS OF THE RECIPROCAL OF A SERIES

By L. Carlitz

1. Introduction. Consider the elliptic function $\wp(u)$ with invariants $g_2 = 4$, $g_3 = 0$, so that $\wp(u)$ satisfies the differential equation

$$\wp'^2(u) = 4\wp^3(u) - 4\wp(u). \tag{1.1}$$

Put

$$\wp(u) = \frac{1}{u^2} + \sum_{m=1}^{\infty} \frac{2^{4m} E_m}{4m} \frac{u^{4m-2}}{(4m-2)!}, \tag{1.2}$$

where E_m are rational. Then Hurwitz[1] has proved the following theorem:

$$E_m = G_m + \frac{1}{2} + \sum \frac{(2a)^{\frac{4m}{p-1}}}{p}, \tag{1.3}$$

where G_m is integral and the summation is extended over those primes $p = 4k + 1$ such that $p - 1 \mid 4m$; furthermore the odd integer a is determined by means of

$$p = a^2 + b^2, \qquad a \equiv b + 1 \pmod 4.$$

The method of proof depends in particular on the complex multiplication of $\wp(u)$, and Hurwitz suggests that like theorems may hold for the coefficients of those elliptic functions that possess complex multiplication. For the case in which the ratio of the periods is an imaginary cube root of unity, this was indeed proved by Matter.[2]

In the present paper we consider the class of series

$$f(u) = \sum_1^{\infty} \frac{c_m u^m}{m!} \qquad (c_1 = 1), \tag{1.4}$$

where the c_m are integral, and assume that the inverse of $f(u)$ has the form

$$\lambda(u) = \sum_1^{\infty} \frac{e_m u^m}{m},$$

Received August 25, 1941.

[1] A. Hurwitz, *Über die Entwickelungscoefficienten der lemniscatischen Functionen*, Mathematische Annalen, vol. 51(1899), pp. 196–226 = *Mathematische Werke*, Basel, 1933, vol. II, pp. 342–373.

[2] K. Matter, *Die den Bernoulli'schen Zahlen analogen Zahlen im Körper der dritten Einheitswurzeln*, Zürich, 1900; reviewed in Jahrbuch der Fortschritte der Math., vol. 31(1900), pp. 204–206.

with integral ϵ_m. We define the rational numbers β_m by means of

$$\frac{u}{f(u)} = \sum_0^\infty \frac{\beta_m u^m}{m!},$$

and prove by elementary methods that

$$\beta_m = G_m - \sum \frac{1}{p} \epsilon_p^{\frac{m}{p-1}}, \tag{1.5}$$

where G_m is integral and the summation extends over all primes p such that $p - 1 \mid m$. If next we put

$$\frac{u^2}{f^2(u)} = 1 + \sum_2^\infty (m-1)\delta_m \frac{u^m}{m!},$$

and make the additional assumption $c_2 = 0$, then we show that $\beta_m + \delta_m$ is integral. This result together with (1.5) includes Hurwitz's theorem (1.3)—except for the term corresponding to the prime 2. Applications to other series are also indicated. It should be noted that in the proof of the main theorems no use is made of complex multiplication or, for that matter, of elliptic functions; on the other hand, certain general ideas, due to Hurwitz, on series of the form (1.4) are fundamental.

2. **Hurwitz series.** We shall call the series

$$H(u) = \sum_{m=0}^\infty \frac{a_m u^m}{m!}, \tag{2.1}$$

where the a_m are ordinary integers, Hurwitz series, briefly H-series. It is easily verified that the sum or product of two H-series is again an H-series. Similarly, the derivative and the (definite) integral

$$H'(u) = \sum_0^\infty \frac{a_{m+1} u^m}{m!}, \qquad \int_0^u H(u)\,du = \sum_1^\infty \frac{a_{m-1}u^m}{m!}$$

are H-series. If in (2.1) the constant term $a_0 = 0$, we shall call $H(u)$ an H_1-series. Since in this case

$$\frac{1}{(k+1)!} H^{k+1}(u) = \int_0^u \frac{1}{k!} H^k(u) H'(u)\,du,$$

we get Hurwitz's theorem that *if $H(u)$ is an H-series without constant term, then $\frac{1}{k!} H^k(u)$ is an H-series for all $k \geqq 1$.* By the congruence

$$\sum_0^\infty \frac{a_m u^m}{m!} \equiv \sum_0^\infty \frac{b_m u^m}{m!} \pmod{k}$$

we shall understand the system of congruences

$$a_m \equiv b_m \pmod{k} \qquad (m = 0, 1, 2, \cdots).$$

Thus the result stated above may be put in the form

$$H^k(u) \equiv 0 \pmod{k!}, \tag{2.2}$$

provided $H(0) = 0$. We remark that if two H-series are congruent (mod k), then their derivatives (and integrals) are also congruent (mod k).

3. Preliminary results. We shall use the following notation. Let

$$f(u) = \sum_{m=1}^{\infty} \frac{c_m u^m}{m!} \qquad (c_1 = 1), \tag{3.1}$$

where the c_m are integral, be an arbitrary H-series without constant term. If we let $\lambda(u)$ denote the inverse of $f(u)$:

$$\lambda(u) = \sum_{m=1}^{\infty} \frac{e_m u^m}{m!}, \qquad \lambda(f(u)) = u = f(\lambda(u)), \tag{3.2}$$

it follows that the e_m are integral and therefore $\lambda(u)$ is also an H_1-series. We now introduce the following

Hypothesis. *For all* $m \geqq 1$,

$$e_m \equiv 0 \pmod{(m-1)!}. \tag{3.3}$$

As a consequence of (3.3) we may put

$$e_m = (m-1)!\epsilon_m \qquad (e_1 = \epsilon_1 = 1), \tag{3.4}$$

where the ϵ_m are integral, and (3.2) becomes

$$\lambda(u) = \sum_{m=1}^{\infty} \frac{\epsilon_m u^m}{m}. \tag{3.5}$$

For brevity we shall refer to a series of the form (3.5) as an HL-series.

Now put

$$f^{p-1}(u) = \sum_{m=p-1}^{\infty} \frac{d_m u^m}{m!}, \tag{3.6}$$

where p is a prime, and $d_m = d_m^{(p-1)}$ is integral for all m. Since by (2.2)

$$f^{p-1} \cdot f \equiv f^p \equiv 0 \pmod{p},$$

it follows from (3.1) and (3.6) that

$$\sum_i \binom{m}{i} c_i d_{m-i} \equiv 0 \pmod{p} \tag{3.7}$$

for $m \geqq p$. Now clearly by (3.6)

$$d_{p-1} = (p-1)! \equiv -1 \pmod{p}. \tag{3.8}$$

In the next place we show that

$$d_m \equiv 0 \pmod{p} \quad \text{for } p \leqq m \leqq 2p-3. \tag{3.9}$$

Indeed take $m = p + 1$ in (3.7):

$$\binom{p+1}{1} c_1 d_p + \binom{p+1}{2} c_2 d_{p-1} \equiv 0,$$

and therefore $d_p \equiv 0 \pmod{p}$ for $p > 2$. It is now evident how (3.9) can be proved by induction. For if $p + 1 \leqq m \leqq 2p - 2$, then (3.7) becomes

$$\binom{m}{1} c_1 d_{m-1} + \cdots + \binom{m}{m-p+1} c_{m-p+1} d_{p-1} \equiv 0; \tag{3.10}$$

if we observe that the last coefficient $\equiv 0$, (3.9) follows immediately. On the other hand, if in (3.10) we take $m = 2p - 1$, we get

$$\binom{2p-1}{1} c_1 d_{2p-2} + \binom{2p-1}{p} c_p d_{p-1} \equiv 0,$$

and therefore

$$d_{2p-2} \equiv -c_p \pmod{p}. \tag{3.11}$$

We now make use of the hypothesis (3.3). Then by (3.2)

$$u = \lambda(f) = \sum_1^\infty \frac{(m-1)!\epsilon_m}{m!} f^m,$$

which evidently implies

$$u \equiv \sum_1^p \frac{\epsilon_m}{m} f^m \pmod{p}. \tag{3.12}$$

Differentiation of (3.12) leads to

$$1 \equiv \sum_1^p \epsilon_m f^{m-1} f', \tag{3.13}$$

whence by division

$$f' \equiv \sum_1^p \eta_m f^{m-1} \quad (\eta_1 = 1), \tag{3.14}$$

where the η_m are integers (mod p).

Now we have

$$D(f^{p-1}) \equiv -f^{p-2} f' \equiv -f^{p-2} - \eta_2 f^{p-1}$$

as follows by multiplying both members of (3.14) by f^{p-2}. Differentiating again and reducing by means of (3.14) we get

$$D^2(f^{p-1}) \equiv \alpha_0 f^{p-3} + \alpha_1 f^{p-2} + \alpha_2 f^{p-1},$$

and similar congruences for derivatives of higher order. In particular for the $(p - 1)$-th derivative,

$$D^{p-1}(f^{p-1}) \equiv A_0 + A_1 f + \cdots + A_{p-1} f^{p-1}. \tag{3.15}$$

It remains to determine the A_m. Inspection of the constant terms in (3.15) shows that

$$A_0 \equiv (p-1)! \equiv -1.$$

In the next place repeated application of (3.9) gives

$$A_m \equiv 0 \qquad (1 \leqq m \leqq p-2).$$

Finally comparing coefficients of u^{p-1}, we get

$$A_{p-1} \equiv -d_{2p-2} \equiv c_p,$$

the latter congruence following from (3.11). Now substituting in (3.15) we have at once

$$D^{p-1}(f^{p-1}) \equiv -1 + c_p f^{p-1} \pmod{p}. \tag{3.16}$$

We can now determine the d_m in (3.6). Clearly (3.16) implies

$$d_{m+p-1} \equiv c_p d_m \qquad \text{for } m \geqq p-1,$$

and therefore by (3.8)

$$d_{m(p-1)} \equiv -c_p^{m-1} \qquad \text{for } m \geqq 1, \tag{3.17}$$

while $d_m \equiv 0$ if $p - 1 \nmid m$. Note that (3.11) is included in (3.17). As an immediate consequence of (3.17) we have the result

$$f^{p-1} \equiv -\sum_1^\infty \frac{c_p^{m-1} u^{m(p-1)}}{(m(p-1))!} \pmod{p}. \tag{3.18}$$

For the sequel it will be convenient to transform (3.18) slightly. We require

$$c_p \equiv -e_p \pmod{p}. \tag{3.19}$$

(We remark that (3.19) is independent of the hypothesis (3.3). We assume only that f—and therefore also λ—is an H_1-series.)

To prove (3.19) consider $D_0^p \lambda^k$, where the subscript indicates that we put $u = 0$ after differentiation. Then

$$D_0^p(\lambda^{k+1}) = D_0^p(\lambda^k \cdot \lambda) \equiv \lambda_0 D_0^p(\lambda^k) + \lambda_0^k D_0^p \lambda \equiv 0,$$

so that

$$D_0^p(\lambda^k) \equiv 0 \pmod{p} \qquad \text{for } k > 1. \tag{3.20}$$

In the next place

$$D^p\left(\frac{1}{p}\lambda^p\right) = D^{p-1}(\lambda^{p-1}\lambda'),$$

$$D_0^p\left(\frac{1}{p}\lambda^p\right) = \lambda_0' D_0^{p-1}(\lambda^{p-1}) + \cdots + \lambda_0^{p-1} D_0^{p-1}\lambda \equiv -1. \tag{3.21}$$

Since (3.2) implies

$$0 = D_0^p u = D_0^p \sum_1^p \frac{c_m}{m!} \lambda^m(u),$$

it follows from (3.20) and (3.21) that

$$D_0^p \lambda + c_p D_0^p \left(\frac{1}{p!} \lambda^p\right) \equiv 0.$$

This reduces to $e_p + c_p \equiv 0$, so that we have proved (3.19).

If we note that by (3.4) $e_p \equiv (p-1)!\epsilon_p \equiv -\epsilon_p \pmod p$, we may state

THEOREM 1. *Let* $f(u) = \sum_1^\infty \frac{c_m u^m}{m!}$ *be an* H_1*-series, let* $\lambda(u) = \sum_1^\infty \frac{e_m u^m}{m!}$ *be the inverse of* $f(u)$. *If* $e_m = (m-1)!\epsilon_m$, *where* ϵ_m *is integral for all* $m \geqq 1$, *then*

$$f^{p-1}(u) \equiv -\sum_1^\infty \frac{\epsilon_p^{m-1} u^{m(p-1)}}{(m(p-1))!} \pmod p, \tag{3.22}$$

where p *is an arbitrary prime.*

The case $p = 2$ requires some further discussion. According to (3.22) we have

$$f(u) \equiv u + \epsilon_2 \sum_2^\infty \frac{u^m}{m!} \pmod 2. \tag{3.23}$$

For ϵ_2 odd, (3.23) implies

$$f^2 \equiv \left(\sum_1^\infty \frac{u^m}{m!}\right)^2 \equiv 2 \sum_2^\infty \frac{u^m}{m!} \pmod 4,$$

so that

$$\frac{f^2}{2} \equiv \sum_2^\infty \frac{u^m}{m!} \equiv u + f \pmod 2,$$

$$\frac{f^3}{2} \equiv uf + f^2 \equiv uf,$$

and therefore

$$\frac{f^3}{2} \equiv \sum_2^\infty \frac{mu^m}{m!} \equiv \sum_2^\infty \frac{u^{2m+1}}{(2m+1)!} \pmod 2. \tag{3.24}$$

For ϵ_2 even we get

$$\frac{f^3}{2} \equiv \frac{u^3}{2} \pmod 2. \tag{3.25}$$

In place of (3.24) and (3.25) we may write the single formula

$$\frac{f^3}{2} \equiv \frac{u^3}{3!} + \epsilon_2 \sum_2^\infty \frac{u^{2m+1}}{(2m+1)!} \pmod 2. \tag{3.26}$$

Note that the right member of (3.26) contains only terms of odd degree.

4. **The main theorems.** Define β_m by means of

$$\frac{u}{f(u)} = \sum_0^\infty \frac{\beta_m u^m}{m!} \qquad (\beta_0 = 1), \tag{4.1}$$

so that β_m is rational for all $m \geqq 0$. Then we have

$$\frac{u}{f} = \frac{\lambda(f)}{f} = \sum_1^\infty \frac{\epsilon_m}{m} f^{m-1}. \tag{4.2}$$

By (2.2) the coefficients in the expansion of f^{m-1} are multiples of $(m - 1)!$. Next we observe that $(m - 1)!$ is a multiple of m except when $m = 4$ or $m =$ prime. Thus (4.2) becomes

$$\frac{u}{f} = h(u) + \sum_p \frac{\epsilon_p}{p} f^{p-1} + \frac{\epsilon_4}{4} f^3, \tag{4.3}$$

where $h(u)$ is an H-series, and the summation is taken over all primes p (including 2). Comparison with (4.1) shows that the denominator of β_m contains only simple factors. For a more precise result we use (3.22) and (3.26). Note that $\frac{1}{4}f^3$ will contribute to the fractional part of β_m only when m is odd; on the other hand, for fixed p, f^{p-1} will contribute only when $p - 1 \mid m$. We may now state our principal result:

THEOREM 2. *Let $f(u)$ satisfy the hypothesis of Theorem 1, and define β_m by means of* (4.1). *Then for m even*

$$\beta_m = G_m - \sum_{p-1 \mid m} \frac{1}{p} \epsilon_p^{\frac{m}{p-1}}, \tag{4.4}$$

while for m odd

$$\beta_1 = \frac{\epsilon_2}{2}, \qquad \beta_3 = G_3 + \frac{\epsilon_2}{2} + \frac{\epsilon_4}{2},$$
$$\beta_m = G_m + \frac{\epsilon_2}{2} + \frac{\epsilon_2 \epsilon_4}{2} \qquad \textit{for } m > 3, \tag{4.5}$$

where G_m is integral, and the summation in (4.4) *is over all primes (including* 2) *such that $p - 1 \mid m$.*

As a first application let $f(u) = e^u - 1$, so that $\lambda(u) = \log(1 + u)$,

$$e_m = (-1)^{m-1}(m-1)!, \qquad \epsilon_m = (-1)^{m-1},$$

and therefore $\lambda(u)$ is an HL-series. In this case evidently $\beta_1 = -\frac{1}{2}$, $\beta_m = G_m$ for $m > 1$ and odd, while for m even (4.4) implies

$$\beta_m = G_m - \sum_{p-1 \mid m} \frac{1}{p}.$$

This is of course the familiar Staudt-Clausen theorem.

We consider next the coefficients in the expansion of

$$\frac{u^2}{f^2(u)} = \sum_0^\infty \gamma_m \frac{u^m}{m!}. \tag{4.6}$$

We shall assume $c_2 = 0$, which implies $\epsilon_2 = \gamma_1 = 0$. Then (4.6) gives

$$u \int_0^u \left(\frac{1}{f^2} - \frac{1}{u^2}\right) du = \sum_2^\infty \frac{\gamma_m}{m-1} \frac{u^m}{m!} = \sum_2^\infty \delta_m \frac{u^m}{m!}, \tag{4.7}$$

thus defining δ_m for $m \geqq 2$. On the other hand by (3.2) and (3.5) we have

$$u = \sum_1^\infty \frac{\epsilon_m}{m} f^m;$$

differentiating and dividing by f^2 leads to

$$\frac{1}{f^2} = \sum_1^\infty \epsilon_m f^{m-3} f',$$

and therefore

$$\int_0^u \left(\frac{1}{f^2} - \frac{1}{u^2}\right) du = \frac{1}{u} - \frac{1}{f} + \sum_3^\infty \epsilon_m \frac{f^{m-2}}{m-2}. \tag{4.8}$$

Comparing (4.8) with (4.7) we get

$$\sum_2^\infty \delta_m \frac{u^m}{m!} + \sum_0^\infty \beta_m \frac{u^m}{m!} = 1 + u \sum_3^\infty \epsilon_m \frac{f^{m-2}}{m-2}.$$

Since the right member is certainly an H-series, it follows at once that $\delta_m + \beta_m$ is integral. This proves

THEOREM 3. *Let $f(u)$ satisfy the hypothesis of Theorem 1, and suppose in addition that $c_2 = 0$. Define δ_m by means of*

$$\frac{u^2}{f^2} = 1 + \sum_2^m (m-1)\delta_m \frac{u^m}{m!}. \tag{4.9}$$

Then $\delta_m + \beta_m$ is integral for all m.

This theorem together with Theorem 2 determines the fractional part of δ_m. However, since we now have $\epsilon_2 = 0$, the final result is somewhat simpler than that implied by (4.4) and (4.5). In particular the additional assumption $c_4 = 0$ implies $\epsilon_4 = 0$ and therefore by (4.5) we now have both β_m and δ_m integral for odd m, while for m even (4.4) gives

$$\beta_m = G_m - \sum_{p-1 \mid m} \frac{1}{p} \epsilon_p^{\frac{m}{p-1}}, \tag{4.10}$$

the summation now extending over odd primes p. A formula similar to (4.10) also holds for δ_m.

As an immediate corollary of Theorems 2 and 3 we note that

$$k(k^m - 1)\beta_m = \beta_{m,k} \tag{4.11}$$

and

$$k(k^m - 1)\delta_m = \delta_{m,k} \tag{4.12}$$

are both integral for all integral k.

For certain applications it is convenient to weaken the hypothesis (3.3) somewhat. Put

$$\epsilon_m = \frac{e_m}{(m-1)!} = \frac{A}{B},$$

where A and B are relatively prime; we now allow B to contain certain "exceptional" primes. Then clearly Theorem 2 still holds, where now G_m is a fraction whose denominator contains only exceptional primes. The same remark applies to Theorem 3.

5. Further results. Let k be a fixed positive integer and assume $f(u)$ of the form

$$f(u) = \sum_0^\infty c_m \frac{u^{km+1}}{(km+1)!} \qquad (c_0 = 1), \tag{5.1}$$

where the c_m are integral (mod k). Then the inverse of $f(u)$ is of the form

$$\lambda(u) = \sum_0^\infty e_m \frac{u^{km+1}}{(km+1)!} \qquad (e_0 = 1), \tag{5.2}$$

where the e_m also are integral (mod k). We now do not assume the hypothesis (3.3). Put

$$\frac{u}{f(u)} = \sum_0^\infty \beta_m \frac{u^{km}}{(km)!} \qquad (\beta_0 = 1), \tag{5.3}$$

then we shall show that

$$\beta_m \equiv e_m \pmod{k} \tag{5.4}$$

for all m.

Now by (5.1) we have

$$f = u + uf_1, \quad \text{where } f \equiv \sum_1^\infty c_m \frac{u^{km}}{(km)!} \pmod{k}.$$

Let $p^s \mid k$; then

$$\frac{f^{p^s}}{p^s!} = \frac{u^{p^s}}{p^s!}(1 + f_1)^{p^s} \equiv \frac{u^{p^s}}{p^s!} \pmod{p^s}.$$

More generally for $m \geqq 1$,

$$\frac{f^{p^s m}}{(p^s m)!} = \frac{u^{p^s m}}{(p^s m)!}(1 + f_1)^{p^s m} \equiv \frac{u^{p^s m}}{(p^s m)!} \pmod{p^s}$$

and therefore if p runs through the prime divisors of k we get

$$\frac{f^{km}}{(km)!} \equiv \frac{u^{km}}{(km)!} \pmod{k}. \tag{5.5}$$

In the next place by (5.2)

$$\frac{u}{f} = \frac{\lambda(f)}{f} = \sum_0^\infty \frac{e_m}{km+1}\frac{f^{km}}{(km)!} \equiv \sum_0^\infty e_m \frac{f^{km}}{(km)!},$$

so that by (5.5)

$$\frac{u}{f} \equiv \sum_0^\infty e_m \frac{u^{km}}{(km)!} \pmod{k}.$$

Comparison with (5.3) leads at once to (5.4). This proves

THEOREM 4. *If $f(u)$ and $\lambda(u)$ are of the form* (5.1) *and* (5.2) *respectively, where c_m and e_m are integral* (mod k), *then for β_m defined by* (5.3) *we have $\beta_m \equiv e_m$* (mod k). *Here k is an arbitrary integer not less than* 1, *and the hypothesis* (3.3) *is not assumed.*

6. **Application to the lemniscate coefficients.** Consider now the special elliptic function $\varphi(u)$ defined by

$$\varphi'^2(u) = 1 - \varphi^4(u), \qquad \varphi(0) = 1. \tag{6.1}$$

As Hurwitz remarks, $\varphi(u)$ is the function used by Eisenstein[3] in his work on the biquadratic reciprocity theorem. It follows from (6.1) that the inverse of $\varphi(u)$ is

$$\lambda(u) = \int_0^u \frac{du}{\sqrt{1-u^4}} = \sum_0^\infty \binom{2m}{m}\frac{u^{4m+1}}{2^{2m}(4m+1)}. \tag{6.2}$$

Except for the power of 2 in the denominator $\lambda(u)$ is an HL-series, as defined in §3; the coefficients ϵ_m are given by

$$\epsilon_m = \begin{cases} \dfrac{1}{2^{2t}}\dbinom{2t}{t} & \text{for } m = 4t+1, \\ 0 & \text{otherwise;} \end{cases} \tag{6.3}$$

in particular note that $\epsilon_2 = \epsilon_4 = 0$. Hurwitz puts

$$\frac{u}{\varphi(u)} = \sum_0^\infty \frac{F_m u^{4m}}{(4m)!}, \tag{6.4}$$

so that F_m is our β_{4m}—we remark that in this case $\beta_m = 0$ for $4 \nmid m$.

[3] G. Eisenstein, *Beiträge zur Theorie der elliptischen Funktionen* I, Journal für die reine und angewandte Mathematik, vol. 30(1846), pp. 185–210 = Mathematische Abhandlungen, Berlin, 1847, pp. 129–154.

We shall first apply Theorem 4. Replace $\lambda(u)$ by $2^{\frac{1}{2}}\lambda(u/2^{\frac{1}{2}})$; then for this function

$$e_{4m+1} = (4m)!\,\epsilon_{4m+1} = \frac{(4m)!}{2^{4m}}\binom{2m}{m},$$

which is easily seen to be odd—indeed $e_{4m+1} \equiv (-1)^m \pmod 4$. Thus we get

$$F_m = 2^{2m}F'_m, \qquad F'_m \equiv (-1)^m \pmod 4. \tag{6.5}$$

We next apply Theorem 2 (see remark at end of §4). We have immediately

$$F_m = \beta_{4m} = G_{4m} - \sum_{p-1|4m} \frac{1}{p}\,\epsilon_p^{\frac{4m}{p-1}}, \tag{6.6}$$

where the summation is over primes of the form $4k + 1$ only, ϵ_p is given by (6.3) and G_{4m} is integral. The result may be improved by using a theorem of Gauss:[4]

$$\epsilon_p = \frac{1}{2^{2k}}\binom{2k}{k} \equiv \frac{3\cdot 7\cdot 11 \cdots (4k-1)}{1\cdot 5\cdot 9 \cdots (4k-3)} \equiv 2a \pmod p, \tag{6.7}$$

where $p = 4k + 1$, and the odd integer a is determined by

$$p = a^2 + b^2, \qquad a \equiv b + 1 \pmod 4.$$

Thus (6.6) becomes

$$F_m = G_{4m} - \sum_{p-1|4m} \frac{1}{p}\,(2a)^{\frac{4m}{p-1}}. \tag{6.8}$$

Turning next to the Weierstrass function $\wp(u) = 1/\varphi^2(u)$ discussed in §1, Theorem 3 may be applied to give

$$E_m = G'_m + \sum_{p-1|4m} \frac{1}{p}\,(2a)^{\frac{4m}{p-1}}, \tag{6.9}$$

where G'_m is a fraction whose denominator is a power of 2. (To determine the fractional part of G'_m most easily one may make use of Hurwitz's formula[5]

$$F_m = (1 + i)^{4m}\{(1 + i)^{4m} - 2\}E_m,$$

whence by (6.5), $G'_m = G_m + \frac{1}{2}$, where G_m is integral, and (6.9) reduces to (1.3). Incidentally, we may go directly from (6.8) and (6.5) to (6.9) by means of the Hurwitz formula.)

7. Other applications. It is clear from §3 that if

$$\lambda'(u) = \sum_1^\infty \epsilon_m u^{m-1} \qquad (\epsilon_1 = 1), \tag{7.1}$$

[4] K. F. Gauss, *Theorie residuorum biquadraticorum*, Werke, vol. 2, p. 90; see also P. Bachmann, *Die Lehre von der kreisteilung*, Leipzig-Berlin, 1921, p. 137.

[5] Hurwitz, loc. cit., formula (21).

where the ϵ_m are arbitrary integers, then Theorem 2 applies; more generally the ϵ_m may be fractions whose denominators are made up of certain "exceptional" primes. If in (7.1) $\epsilon_2 = 0$, then Theorem 3 also applies; if $\lambda(u)$ is odd, say, then both Theorem 3 and Theorem 4 ($k = 2$) may be applied. These conditions are of course all satisfied in the lemniscate case. Various generalizations are immediate. For example, the theorems evidently apply to the hyperelliptic case:

$$\lambda'(u) = \{(1 - u^2)(1 - \alpha_1 u^2) \cdots (1 - \alpha_k u^2)\}^{-\frac{1}{2}}, \tag{7.2}$$

where the α_i are integers (or even rational, in which case the prime factors of the g.c.d. of the α_i may be exceptional). The case

$$\lambda'(u) = (1 - u^6)^{-\frac{1}{2}} \tag{7.3}$$

is that treated by Matter.[6] Put

$$u^2 \wp(u) = \frac{u^2}{\varphi^2(u)} = 1 + \sum_1^\infty \frac{2^{6m} E_m}{6m} \frac{u^{6m}}{(6m - 2)!},$$

where as usual $\varphi(u)$ is the inverse of the function defined by (7.3). Then Theorem 3 gives immediately

$$E_m = G'_m + \sum_{p-1 \mid 6m} \frac{1}{p} \epsilon_p^{\frac{6m}{p-1}}, \tag{7.4}$$

the summation extending over primes p of the form $6k + 1$ such that $p - 1 \mid 6m$; G'_m is a fraction whose denominator is a power of 2. (According to Matter $G'_m = G_m + \dfrac{(-1)^m}{4}$, G_m integral.) The coefficient ϵ_p is determined by

$$\epsilon_m = \begin{cases} \dfrac{1}{2^{2t}} \dbinom{2t}{t} & \text{for } m = 6t + 1, \\ 0 & \text{otherwise}; \end{cases}$$

and by means of a formula[7] similar to (6.7) we have $\epsilon_p \equiv 2\alpha \pmod p$, where $p = \alpha^2 + 3\beta^2$.

More generally the last result (7.4) may be extended to the case

$$\lambda'(u) = (1 - u^k)^{s/t},$$

where s and t are relatively prime; the primes dividing t may be exceptional—if, however, k is a multiple of t, then Theorem 4 applies and the denominator of β_m contains only primes $p \equiv 1 \pmod k$. In like manner it is evident how (7.2) may be generalized; however, in this case there is usually no simple explicit formula for ϵ_p. It would be of interest to know when a formula like (6.7) is available for these generalizations.

Duke University.

[6] See footnote 2.

[7] See Bachmann, loc. cit., p. 141.

CERTAIN QUANTITIES TRANSCENDENTAL OVER $GF(p^n, x)$

By L. I. Wade

1. **Introduction.** Let $GF(q)$, $q = p^n$, denote a fixed finite, Galois, field of order q; let x be an indeterminate over the field $GF(q)$. If x is adjoined to the field $GF(q)$, a new field $GF(q, x)$ is obtained. We are interested here in the nature of certain quantities over the field $GF(q, x)$. "Transcendental" throughout this paper will mean "transcendental over $GF(q, x)$".

Certain polynomials[1] in $GF(q)$ and certain functions connected with the polynomials in $GF(q)$ are of particular interest. Place

$$
\begin{aligned}
[k] &= x^{q^k} - x, \\
F_k &= [k][k-1]^q \cdots [1]^{q^{k-1}}, \\
L_k &= [k][k-1] \cdots [1], \\
F_0 &= L_0 = 1.
\end{aligned}
$$

If $\psi_k(t) = \prod (t - E)$, extended over all polynomials $E(x)$ of $GF(q)$ of degree $<k$, where k is an arbitrary positive integer, then

$$\psi_k(t) = \sum_{j=0}^{k} (-1)^j \frac{F_k}{F_j L_{k-j}^{q^j}} t^{q^j}.$$

The function

$$\psi(t) = \sum_{k=0}^{\infty} \frac{(-1)^k t^{q^k}}{F_k} \tag{1.1}$$

has the property that

$$\psi(E\xi) = 0 \tag{1.2}$$

for all polynomials E in $GF(q)$ and for a fixed

$$\xi = \lim_{k\to\infty} \frac{[1]^{q^k/(q-1)}}{L_k}. \tag{1.3}$$

Also, for a polynomial M of degree m,

$$\psi(Mt) = \sum_{j=0}^{m} \frac{(-1)^j}{F_j} \psi_j(M)\psi^{q^j}(t). \tag{1.4}$$

Received August 27, 1941.

[1] For the properties of the polynomials and functions stated below see L. Carlitz, *On certain functions connected with polynomials in a Galois field*, Duke Mathematical Journal, vol. 1(1935), pp. 137–168. Other references are given there.

If $\lambda(t)$ denotes the inverse function of $\psi(t)$, i.e., $\psi(\lambda(t)) \equiv t$, then

$$\lambda(t) = \sum_{j=0}^{\infty} \frac{t^{q^j}}{L_j}. \tag{1.5}$$

As we shall see it is very easy to prove that $\psi(1)$ is transcendental.[2] By making use of property (1.2) of $\psi(t)$, we prove that the quantity ξ is transcendental over $GF(q, x)$. More generally, $\psi(\alpha)$ is transcendental for $\alpha \neq 0$ algebraic. From this it follows that $\lambda(\alpha)$ is transcendental for $\alpha \neq 0$ algebraic.

If E is a polynomial, the theorem for $\psi(\alpha)$ includes the transcendence of

$$\sum_{j=0}^{\infty} \frac{(-1)^j E^{q^j}}{F_j},$$

but it does not even include the slightly modified series[3]

$$\sum_{j=0}^{\infty} c_j \frac{E^{q^j}}{F_j} \qquad (c_j \neq 0 \text{ in } GF(q)). \tag{1.6}$$

It does not seem possible to generalize the theorem for $\psi(\alpha)$ without a totally different method of proof since we require the use of the multiplication theorem (1.4). However it is easy to prove the transcendence of (1.6). This suggests a consideration of series of the form

$$\sum_{j=0}^{\infty} \frac{B_j}{F_j}, \tag{1.7}$$

where the B_j are polynomials with an infinite number of them not zero. In this direction, we prove that (1.7) is transcendental if degree $B_k \leqq (q-1)(k-1)q^{k-1} - b_k q^k$, where $\lim_{k\to\infty} b_k = \infty$. For reasons we shall give below, it does not seem likely that we can improve this theorem with our methods.

A proof is also given of the transcendence of the interesting series

$$\sum_{k=1}^{\infty} \frac{F_{k-1}^q}{F_k} = \sum_{k=1}^{\infty} \frac{1}{[k]}.$$

For brevity we shall refer to this as the bracket series.

The writer wishes to express his gratitude to Professor Carlitz, who suggested the problem and who offered many suggestions throughout the preparation of the paper.

2. Notation and a fundamental lemma.

We shall use the expressions *is integral* and *is a polynomial* (in $GF(q)$) interchangeably. The abbreviation deg will be used for degree. We note that

[2] $\psi(1)$ is in some ways an analogue of e and hence $\psi(t)$ is an analogue of e^t. ξ becomes an analogue of π or rather πi and $\lambda(t)$ of log t.

[3] Unless $q = 2$ and the two series are identical.

$$\deg\,[k] = q^k;$$
$$\deg F_k = kq^k;$$
$$\deg L_k = (q^{k+1} - q)/(q - 1).$$

Also define

$$[k, d] = [k][k-1]^q \cdots [d]^{q^{k-d}} \qquad (d \leqq k)$$
$$= \frac{F_k}{F_{d-1}^{q^{k-d+1}}}.$$

If D and E are defined as the products of brackets $[k]$, we shall denote by $\mathcal{E}(j, D/E)$ the number of times $[j]$ occurs in D by definition minus the number of times it occurs in E. All divisibility properties are disregarded. For example, if $j \leqq k$,

$$\mathcal{E}(j, F_k) = q^{k-j}; \qquad \mathcal{E}(j, F_k/L_k) = q^{k-j} - 1; \qquad \mathcal{E}(k, L_{2k}) = 1.$$

The following fundamental lemma[4] allows us to make full use of the characteristic p.

LEMMA 2.1. *Every polynomial in $GF(q, x)$ divides a linear polynomial*

$$\sum_{j=l}^{m} A_j t^{q^j} \qquad (A_l \neq 0, A_m \neq 0) \tag{2.1}$$

where the A_j are integral.

Proof. Let $f(t)$ of degree m be the given polynomial. Divide all the q-th powers of t by $f(t)$;

$$t^{q^i} \equiv C_{m-1}^{(i)} t^{m-1} + \cdots + C_0^{(i)} \quad (\mathrm{mod}\, f(t)) \qquad (i = 0, 1, 2, \cdots),$$

where the $C_j^{(i)}$ are in $GF(q, x)$. The powers $1, t, t^2, \cdots$ on the right side of the first $v \leqq m$ congruences can be eliminated and on the left side we obtain a linear polynomial. Multiplication by a suitable polynomial (in $GF(q)$) will make all of the coefficients of the linear polynomial integral and the proof is complete.

3. $\psi(1)$ and related series.

Although the transcendence of $\psi(1)$ follows from later theorems, we shall give a separate proof because of its simplicity and because it illustrates the general method to be followed. Suppose that $\psi(1)$ is algebraic. By Lemma 2.1, we may suppose that $\psi(1)$ is a root of the linear polynomial (2.1). Hence, we have

$$\sum_{j=l}^{m} A_j \sum_{k=0}^{\infty} \frac{(-1)^k}{F_k^{q^j}} = \sum_{k=0}^{\infty} \frac{D_k}{F_k} = 0, \tag{3.1}$$

[4] See O. Ore, *A special class of polynomials*, Transactions of the American Mathematical Society, vol. 35(1933), pp. 559–584.

where

$$D_k = \sum_{i+j=k} \frac{(-1)^i A_j F_k}{F_i^{q^j}} \qquad (j = l, \cdots, m).$$

Place

$$(3.2) \qquad I = F_\beta \sum_{k=0}^{\beta} \frac{D_k}{F_k};$$
$$Q = F_\beta \sum_{k=\beta+1}^{\infty} \frac{D_k}{F_k}.$$

β will be chosen later. Therefore, (3.1) is

$$(3.3) \qquad I + Q = 0 \qquad (\text{all } \beta).$$

I is obviously integral by the definition of F_k. Furthermore,

$$\begin{aligned} I &\equiv D_\beta \qquad (\text{mod } [\beta - l]) \\ &\equiv (-1)^{\beta-l} A_l [\beta] \cdots [\beta - l + 1]^{q^{l-1}} \\ &\equiv (-1)^{\beta-l} A_l F_l \not\equiv 0 \end{aligned}$$

for β sufficiently large. Therefore, for all sufficiently large β, $I \neq 0$. On the other hand, let N be any term of Q. Note that

$$\begin{aligned} \deg D_\beta &\leq \max_j (jq^\beta + \deg A_j) \\ &= mq^\beta + \deg A_m \end{aligned}$$

for β sufficiently large. Hence

$$\begin{aligned} \deg N &\leq \beta q^\beta - (\beta + 1)q^{\beta+1} + mq^{\beta+1} + \deg A_m \\ &\to -\infty, \qquad \beta \to \infty. \end{aligned}$$

Therefore, we may choose β so large that every term of Q is of negative degree, and $I \neq 0$ is integral. This contradicts (3.3) and we have the

THEOREM 3.1. *$\psi(1)$ is transcendental.*

Now let us consider a set of series that includes $\psi(1)$. We shall prove the following

THEOREM 3.2. *If $B_0, B_1, \cdots$ satisfy the three conditions*
(i) the B_k are polynomials,
(ii) an infinite number of the B_k are not zero,
(iii) $\deg B_k \leq (q - 1)(k - 1)q^{k-1} - b_k q^k$ for all k sufficiently large, where $b_k \to \infty$ as $k \to \infty$, then the quantity

$$(3.4) \qquad \sum_{k=0}^{\infty} \frac{B_k}{F_k}$$

is transcendental.

Proof. Suppose that (3.4) is algebraic. By Lemma 2.1, we may suppose that (3.4) is a root of the linear polynomial (2.1). We have

$$\sum_{j=l}^{m} A_j \sum_{k=0}^{\infty} \frac{B_k^{q^j}}{F_k^{q^j}} = \sum_{k=0}^{\infty} \frac{D_k}{F_k} = 0, \tag{3.5}$$

where

$$D_k = \sum_{i+j=k} \frac{A_j B_i^{q^j} F_k}{F_i^{q^j}} \qquad (j = l, \cdots, m). \tag{3.6}$$

Define I and Q by (3.2), where the D_k are those of (3.6). β will be chosen later. Then by (3.5)

$$I + Q = 0 \qquad (\text{all } \beta). \tag{3.7}$$

We note that I is integral since the B_k are by (i). Also, using (iii), we have

$$\begin{aligned} \deg D_k &\leqq \max_j (q^j \deg B_i + jq^k + \deg A_j) \\ &\leqq \max_j ((q-1)(i-1)q^{k-1} - b_i q^k + \deg A_j + jq^k) \\ &\leqq (q-1)(k-1)q^{k-1} + mq^k - b_k' q^k \end{aligned}$$

for all k sufficiently large, where

$$b_k' = \min_{k \geqq i \geqq k-m} b_i .$$

Let N be any term of Q. Therefore,

$$\begin{aligned} \deg N &\leqq \beta q^\beta - (\beta+1)q^{\beta+1} + (q-1)\beta q^\beta + mq^{\beta+1} - b_{\beta+1}' q^{\beta+1} \\ &= -q^{\beta+1} + mq^{\beta+1} - b_{\beta+1}' q^{\beta+1} \\ &\to -\infty, \qquad \beta \to \infty. \end{aligned}$$

Hence, for β sufficiently large, every term of Q is of negative degree. Therefore, since I is integral, (3.7) implies

$$I = 0, \qquad Q = 0 \tag{3.8}$$

for all β sufficiently large.

Now

$$I \equiv D_\beta \qquad (\text{mod } F_\beta / F_{\beta-1}). \tag{3.9}$$

Also

$$\begin{aligned} \deg F_\beta/F_{\beta-1} - \deg D_\beta &\geqq \beta q^\beta - (\beta-1)q^{\beta-1} - (q-1)(\beta-1)q^{\beta-1} - mq^\beta + b_\beta' q^\beta \\ &\geqq q^\beta(1 - m + b_\beta') \\ &\to \infty, \qquad \beta \to \infty. \end{aligned}$$

Therefore, (3.8) and (3.9) imply

$$D_\beta = 0 \qquad \text{(all } \beta \text{ sufficiently large).} \tag{3.10}$$

Hence there exists an $\alpha > \deg A_l$ such that for all $k \geqq \alpha + l$, $D_k = 0$. We see that (by the definition of F_k and (3.6))

$$\frac{F_\alpha^{q^l}}{F_{\alpha+l}} D_{\alpha+l} \equiv A_l B_\alpha^{q^l} \qquad (\text{mod } [\alpha]^{q^l}).$$

Hence $[\alpha]^{q^l}$ divides $A_l B_\alpha^{q^l}$ since $D_{\alpha+l} = 0$. We proceed by induction. Suppose that $[k, \alpha]^{q^l}$ divides $A_l^{(q^{k-\alpha+1}-1)/(q-1)} B_k^{q^l}$ for $\beta - 1 \geqq k \geqq \alpha$. Then,

$$\begin{aligned} A_l^{q(q^{\beta-\alpha}-1)/(q-1)} \frac{F_\beta^{q^l}}{F_{\beta+l}} D_{\beta+l} &= A_l^{q(q^{\beta-\alpha}-1)/(q-1)} \sum_{j=l}^{m} \frac{A_j B_{\beta+l-j}^{q^j} F_\beta^{q^l}}{F_{\beta+l-j}^{q^j}} \\ &\equiv A_l^{(q^{\beta-\alpha+1}-1)/(q-1)} B_\beta^{q^l} \qquad (\text{mod } [\beta, \alpha]^{q^l}). \end{aligned}$$

Therefore, $[\beta, \alpha]^{q^l}$ divides $A_l^{(q^{\beta-\alpha+1}-1)/(q-1)} B_\beta^{q^l}$ for all $\beta \geqq \alpha$; and B_β is either zero or

$$\begin{aligned} q^l \deg B_\beta &\geqq \deg [\beta, \alpha]^{q^l} - \deg A_l \cdot (q^{\beta-\alpha+1} - 1)/(q - 1) \\ &\geqq (\beta - \alpha + 1)q^{\beta+l} - q^{\beta-\alpha+1} \cdot \deg A_l \\ &\geqq (\beta - \alpha)q^{\beta+l} \\ &> (q - 1)(\beta - 1)q^{\beta-1+l} - b_\beta q^{\beta+l} \end{aligned}$$

for β sufficiently large. This is a contradiction of (ii) and (iii) and the proof is complete.

Theorem 3.2 seems fairly close to the best we can do in this direction (that is, with this method of proof). The only natural multiplier for the series (3.4) is F_β. However, if we allow b_k to be some fixed constant, we can give series which, after multiplication by F_β, have terms that are not integral although they become arbitrarily large in degree as β increases. Consider the case $b_k = 0$ $(k = 0, 1, \cdots)$. F_{k-1}^{q-1} is of degree $(q - 1)(k - 1)q^{k-1}$, and

$$\sum_{k=1}^{\infty} \frac{F_{k-1}^{q-1}}{F_k} = \sum_{k=1}^{\infty} \frac{1}{[k]F_{k-1}} \tag{3.11}$$

would satisfy all of the conditions. If we suppose that the series is a root of (2.1) and multiply by F_β, we have

$$F_\beta \sum_{j=l}^{m} A_j \sum_{k=1}^{\infty} \frac{1}{[k]^{q^j} F_{k-1}^{q^j}} = 0.$$

The term

$$N = \frac{F_\beta A_m}{[\beta + 1 - m]^{q^m} F_{\beta-m}^{q^m}}$$

is not integral if β is large enough, but

$$\begin{aligned}\deg N &= \beta q^\beta - (\beta - m)q^\beta - q^{\beta+1} + \deg A_m \\ &= mq^\beta - q^{\beta+1} + \deg A_m \\ &\to \infty, \qquad \beta \to \infty\end{aligned}$$

if $m > q$.

However, we can prove that the quantity (3.11) is transcendental if we multiply by $L_{\beta-l}^{q^m} F_{\beta-l-1}^{q^l}$. This makes it clear that the choice of multiplier is important. In special cases, there is no general method of obtaining the proper multiplier if it exists.

4. **The bracket series.** Closely related to the series (3.11) is the interesting series

$$\sum_{k=1}^{\infty} \frac{F_{k-1}^q}{F_k} = \sum_{k=1}^{\infty} \frac{1}{[k]}, \tag{4.1}$$

which we shall prove transcendental in this section. Since

$$\frac{[2\beta]^{q^m}}{[2\beta + s]^{q^j}} = [2\beta + s]^{q^{m-s} - q^j} - \frac{[s]^{q^{m-s}}}{[2\beta + s]^{q^j}} \qquad (s = 1, \cdots, m - j),$$

we can write

$$\frac{L_{2\beta}^{q^m}}{[2\beta + s]^{q^j}} = [2\beta + s]^{q^{m-s} - q^j} L_{2\beta-1}^{q^m} - \frac{[s]^{q^{m-s}} L_{2\beta-1}^{q^m}}{[2\beta + s]^{q^j}}.$$

By repetition of this process, we can easily prove by induction the

LEMMA 4.1. *For $s = 1, \cdots, m - j$,*

$$\begin{aligned}\frac{L_{2\beta}^{q^m}}{[2\beta + s]^{q^j}} = [2\beta + s]^{q^{m-s} - q^j} L_{2\beta-1}^{q^m} + \sum_{k=0}^{m-s-j-1} (-1)^{k+1}[s]^{q^{m-s}}[s + 1]^{q^{m-s-1}} \cdots \\ [s + k]^{q^{m-s-k}}[2\beta + s]^{q^{m-s-k-1} - q^j} L_{2\beta-k-2}^{q^m} \\ + \frac{(-1)^{m-s-j-1}[s]^{q^{m-s}} \cdots [m - j]^{q^j} L_{2\beta-m+s+j-1}^{q^m}}{[2\beta + s]^{q^j}}.\end{aligned} \tag{4.2}$$

Suppose that (4.1) is algebraic—a root of (2.1). We have[5]

$$\frac{L_{2\beta}^{q^m}}{L_\beta} \sum_{i=l}^{m} A_i \sum_{k=1}^{\infty} \frac{1}{[k]^{q^i}} = 0 \qquad \text{(all } \beta). \tag{4.3}$$

Since $[k]$ divides $[2k] = [k]^{q^k} + [k]$,

$$\frac{L_{2\beta}^{q^m}}{L_\beta} \sum_{i=l}^{m} A_i \sum_{k=1}^{2\beta} \frac{1}{[k]^{q^i}} \tag{4.4}$$

is integral.

[5] $L_\beta^{q^m}$ would suffice as a multiplier if $q > 2$.

Using (4.2), we can write (4.3) in two parts, i.e.,

$$I + Q = 0 \qquad \text{(all } \beta\text{)}, \tag{4.5}$$

where

$$I = \frac{L_{2\beta}^{q^m}}{L_\beta} \sum_{j=l}^{m} A_j \sum_{k=1}^{2\beta} \frac{1}{[k]^{q^j}} + \sum_{j=l}^{m} A_j \sum_{s=1}^{m-j} [2\beta + s]^{q^{m-s}-q^j} \frac{L_{2\beta-1}^{q^m}}{L_\beta}$$
$$+ \sum_{j=l}^{m} A_j \sum_{s=1}^{m-j} \sum_{k=0}^{m-s-j-1} (-1)^{k+1} \prod_{i=0}^{k} [s+i]^{q^{m-s-i}} \cdot [2\beta + s]^{q^{m-s-k}} \frac{L_{2\beta-k-2}^{q^m}}{L_\beta},$$

and

$$Q = \sum_{j=l}^{m} A_j \sum_{s=1}^{m-j} (-1)^{m-s-j+1} \prod_{i=0}^{m-s-j} [s+i]^{q^{m-s-i}} \frac{L_{2\beta-m+s+j-1}^{q^m}}{[2\beta+s]^{q^j} L_\beta}$$
$$+ \frac{L_{2\beta}^{q^m}}{L_\beta} \sum_{j=l}^{m} A_j \sum_{k=2\beta+m+1-j}^{\infty} \frac{1}{[k]^{q^j}}.$$

Since (4.4) is integral, I is obviously integral for $\beta > m$. With β sufficiently large $I \neq 0$, for the minimum value of $2\beta - k - 2$ is easily seen to be $\geqq 2\beta - m$. Hence

$$I \equiv A_m L_{2\beta-m-1}^{q^m} L_m^{q^m} \qquad \pmod{[2\beta - m]} \tag{4.6}$$
$$\not\equiv 0$$

for β sufficiently large. Since for $s + j > m + t$

$$\deg \frac{L_{2\beta+t}^{q^m}}{L_\beta [2\beta+s]^{q^j}} = \frac{q^{2\beta+m+t+1} - q^{m+1}}{q-1} - q^{2\beta+s+j} - \frac{q^{\beta+1} - q}{q-1}$$
$$\to -\infty, \qquad \beta \to \infty,$$

every term of Q is of negative degree for β sufficiently large. This, with $I \neq 0$, implies

$$I + Q \neq 0,$$

a contradiction of (4.5). We therefore have[6]

THEOREM 4.1. *The series*

$$\sum_{k=1}^{\infty} \frac{1}{[k]}$$

is transcendental.

5. **A set of lemmas.** In this section we prove a set of lemmas to be used in subsequent sections.

[6] The significance of the polynomials $[k]$ is well known. See for example L. E. Dickson, *Linear Groups*, 1901, and Lemma 5.8 below.

LEMMA 5.1. *If*

$$q^{k_0} + \cdots + q^{k_r} \leqq q^{\beta} + \cdots + q^{\beta-r} \qquad (k_0 \geqq \cdots \geqq k_r) \tag{5.1}$$

and $k_i \geqq \beta - l$, *then*

$$q^{k_0} + \cdots + q^{k_i} \leqq q^{\beta} + \cdots + q^{\beta-l}.$$

Proof. Write

$$q^{k_0} + \cdots + q^{k_i} = \delta_0 q^{\beta-s_0} + \cdots + \delta_j q^{\beta-s_j},$$

where $0 < \delta_j \leqq q - 1, 0 \leqq s_0 < s_1 < \cdots < s_j \leqq l$. If $s_v = v$ $(v = 0, 1, \cdots, j)$, then $j \leqq l$ since $s_j \leqq l$. Also δ_i $(i = 0, \cdots, j)$ must be one for (5.1) to hold (since $2q^k > q^k + q^{k-1} + \cdots + 1$). The lemma follows in this case.

In the other case, let v be the first subscript for which $s_v > v$. Since $s_v \leqq l$, $v < l$. As before $\delta_0 = \delta_1 = \cdots = \delta_{v-1} = 1$. Therefore

$$\begin{aligned} q^{k_0} + \cdots + q^{k_i} &= \delta_0 q^{\beta-s_0} + \cdots + \delta_j q^{\beta-s_j} \\ &\leqq q^{\beta} + \cdots + q^{\beta-v+1} + (q-1)(q^{\beta-v-1} + \cdots + q^{\beta-l}) \\ &= q^{\beta} + \cdots + q^{\beta-v+1} + q^{\beta-v} - q^{l} \\ &< q^{\beta} + \cdots + q^{\beta-v+1} + q^{\beta-v} + \cdots + q^{\beta-l}. \end{aligned}$$

This completes the proof.

LEMMA 5.2. *If*

$$K_\beta = K_\beta(r) = F_\beta F_{\beta-1} \cdots F_{\beta-r} \tag{5.2}$$

and (5.1) *holds, then*

$$K' = \frac{K_\beta}{F_{k_0} \cdots F_{k_r}} \qquad (k_0 \geqq \cdots \geqq k_r) \tag{5.3}$$

is a polynomial.

Proof. Consider any $[j]$ in K'. If $j \leqq \beta - r$, $k_i \geqq j$ and $k_{i+1} < j$,

$$\begin{aligned} \mathfrak{S}(j, K') &= q^{\beta-j} + \cdots + q^{\beta-r-j} - (q^{k_0-j} + \cdots + q^{k_i-j}) \\ &\geqq q^{\beta-j} + \cdots + q^{\beta-r-j} - (q^{k_0-j} + \cdots + q^{k_r-j}) \geqq 0 \end{aligned}$$

by (5.1). If $j = \beta - l, l < r$,

$$\mathfrak{S}(j, K') = q^{\beta-j} + \cdots + q^{\beta-l-j} - (q^{k_0-j} + \cdots + q^{k_i-j}),$$

where $k_i \geqq \beta - l$, $k_{i+1} < \beta - l$, and in this case $\mathfrak{S}(j, K') \geqq 0$ by Lemma 5.1. This completes the proof.

LEMMA 5.3. *If*

$$q^{k_0} + \cdots + q^{k_r} < q^\beta \qquad (k_0 \geqq \cdots \geqq k_r) \tag{5.4}$$

and

$$\frac{r+1}{q-1} \leqq d < \frac{r+1}{q-1} + 1, \tag{5.5}$$

then

$$\begin{aligned} q^{k_0} + \cdots + q^{k_r} &\leqq (q-1)q^{\beta-1} + \cdots + (q-1)q^{\beta-d} \\ &= q^\beta - q^{\beta-d}. \end{aligned}$$

Proof. Write, as before,

$$q^{k_0} + \cdots + q^{k_r} = \delta_0 q^{\beta-s_0} + \cdots + \delta_j q^{\beta-s_j}.$$

s_0 must be $\geqq 1$ for (5.4) to hold and the right side has its greatest value when each $\delta_i = q - 1$, and $s_i = i + 1$ $(i = 0, \cdots, j)$. Since $\delta_0 + \delta_1 + \cdots + \delta_j \leqq r + 1 \leqq d(q-1)$, the result follows.

LEMMA 5.4. *If* (5.4) *holds and if* $k_i \geqq \beta - l$ $(l < d)$, *then*

$$\begin{aligned} q^{k_0} + \cdots + q^{k_i} &\leqq (q-1)q^{\beta-1} + \cdots + (q-1)q^{\beta-l} \\ &= q^\beta - q^{\beta-l}. \end{aligned}$$

Proof. If we write $q^{k_0} + \cdots + q^{k_i} = \delta_0 q^{\beta-s_0} + \cdots + \delta_j q^{\beta-s_j}$, then $s_j \leqq l$, and since (5.4) makes $s_0 \geqq 1$, the result follows.

LEMMA 5.5. *If condition* (5.4) *holds, then* (5.3) *is divisible by*

$$M = \frac{K_\beta}{F_{\beta-1}^{q-1} \cdots F_{\beta-d}^{q-1}}, \tag{5.6}$$

where d is defined by (5.5).

Proof. We must show that

$$\frac{K_\beta}{F_{k_0} \cdots F_{k_r}} \Big/ \frac{K_\beta}{F_{\beta-1}^{q-1} \cdots F_{\beta-d}^{q-1}} = \frac{F_{\beta-1}^{q-1} \cdots F_{\beta-d}^{q-1}}{F_{k_0} \cdots F_{k_r}} = M' \tag{5.7}$$

is integral if $q^{k_0} + \cdots + q^{k_r} < q^\beta$. Consider any $[j]$. If $j \leqq \beta - d$,

$$\begin{aligned} \mathfrak{S}(j, M') &\geqq (q-1)(q^{\beta-1-j} + \cdots + q^{\beta-d-j}) - (q^{k_0-j} + \cdots + q^{k_r-j}) \\ &\geqq 0, \end{aligned}$$

by Lemma 5.3. If $j = \beta - l$, $l < d$,

$$\mathfrak{S}(j, M') = (q-1)(q^{\beta-1-j} + \cdots + q^{\beta-l-j}) - (q^{k_0-j} + \cdots + q^{k_i-j}),$$

where $k_i \geqq \beta - l$, $k_{i+1} < \beta - l$, and in this case $\mathfrak{S}(j, M') \geqq 0$ by Lemma 5.4. This completes the proof.

The following lemma follows from the proofs of the fundamental theorem on symmetric sums.[7]

LEMMA 5.6. *If* $\alpha_0, \cdots, \alpha_r$ *are the roots of the equation*

$$\sum_{j=0}^{r+1} C_{r+1-j} t^j = 0 \qquad (C_0 = 1, C_{r+1} \neq 0) \tag{5.8}$$

where the C_j *are polynomials and*

$$c = \max(\deg C_1, \cdots, \deg C_{r+1}),$$

then the symmetric sum

$$\sum \alpha_0^{q^{k_0}} \cdots \alpha_r^{q^{k_r}} \qquad (k_0 \geqq \cdots \geqq k_r)$$

is a polynomial and is of degree at most cq^{k_0}.

LEMMA 5.7. *If*

$$q^{k_0} + \cdots + q^{k_i} = q^\beta \qquad (k_0 \geqq \cdots \geqq k_r),$$

then $k_i \geqq \beta - d$, *where d is defined by* (5.5).

Proof. Suppose $k_i < \beta - d$. Hence

$$q^{k_0} + \cdots + q^{k_{i-1}} < q^\beta$$

and by Lemma 5.4 $\leqq q^\beta - q^{\beta-d}$, and the result follows immediately.

The following lemma is well known.[8]

LEMMA 5.8. *If E is an irreducible polynomial of degree e, then E divides* $[l]$ *if and only if e divides l, and then only one time.*

LEMMA 5.9. *If*

$$q^{k_0} + \cdots + q^{k_j} \leqq q^\beta + \cdots + q^{\beta-r} \qquad (j < r),$$

then

$$q^{k_0} + \cdots + q^{k_j} \leqq q^\beta + \cdots + q^{\beta-j}.$$

Proof. Write $q^{k_0} + \cdots + q^{k_j} = \delta_0 q^{\beta-s_0} + \cdots + \delta_i q^{\beta-s_i}$, where $0 < \delta_i \leqq q - 1$ and $0 \leqq s_0 < s_1 < \cdots < s_i$. Then $i \leqq j$ and the right side has its maximum value when $\delta_0 = \delta_1 = \cdots = \delta_i = 1$, $s_h = h$ $(h = 0, \cdots, i)$, and $i = j$. The result follows.

LEMMA 5.10. *If*

$$q^\beta < q^{k_0} + \cdots + q^{k_r} \leqq q^\beta + \cdots + q^{\beta-r} \qquad (k_0 \geqq \cdots \geqq k_r),$$

there exists an i such that

$$q^{k_0} + \cdots + q^{k_i} = q^\beta.$$

[7] See the statements of Theorems 68 and 69 in O. Perron, *Algebra*, vol. 1, 1932.

[8] L. E. Dickson, loc. cit.

Proof. Since $k_0 \leqq \beta$, and since the lemma would follow immediately for $k_0 = \beta$, there exists an i such that

$$q^{k_0} + \cdots + q^{k_{i-1}} < q^\beta \leqq q^{k_0} + \cdots + q^{k_i},$$

that is

$$q^{k_0} + \cdots + q^{k_{i-1}} = q^\beta - A, \qquad A > 0.$$

Hence $q^{k_{i-1}}$ divides A and, therefore, q^{k_i} divides A. Also

$$q^{k_0} + \cdots + q^{k_i} = q^\beta + B, \qquad B \geqq 0.$$

Subtracting, we have

$$q^{k_i} = A + B$$

and $A = q^{k_i}$ and $B = 0$. The lemma follows.

6. **Transcendence of ξ.** In this section we prove the following

THEOREM 6.1. *ξ is transcendental.*

Proof. Suppose that ξ is algebraic. We may further suppose that ξ is an algebraic integer, i.e., that it satisfies equation (5.8) with $\xi = \alpha_0$. Otherwise we merely multiply ξ by a suitable polynomial. We may suppose that $r + 1 = q^s$. Write

$$\gamma_i = E\alpha_i \qquad (i = 0, 1, \cdots, r),$$

where E is an *irreducible* polynomial of degree $e > d$, with d defined by (5.5), and $> \deg C_{r+1}$. Place

$$S(k_0, \cdots, k_r) = (-1)^{k_0+\cdots+k_r} \sum \gamma_0^{q^{k_0}} \cdots \gamma_r^{q^{k_r}},$$
$$\bar{S}(k_0, \cdots, k_r) = (-1)^{k_0+\cdots+k_r} \sum \alpha_0^{q^{k_0}} \cdots \alpha_r^{q^{k_r}},$$

where the sums are symmetric sums of the quantities involved. Then

$$S(k_0, \cdots, k_r) = E^{q^{k_0}+\cdots+q^{k_r}} \bar{S}(k_0, \cdots, k_r). \tag{6.1}$$

From (1.2),

$$\psi(\xi) = \psi(\alpha_0) = \psi(\gamma_0) = 0.$$

Hence

$$\begin{aligned} K_\beta \prod_i \psi(\gamma_i) &= K_\beta \prod_i \sum_{k=0}^{\infty} \frac{(-1)^k \gamma_i^{q^k}}{F_k} \\ &= K_\beta \sum_{k_0 \geqq \cdots \geqq k_r} \frac{S(k_0, \cdots, k_r)}{F_{k_0} \cdots F_{k_r}} \\ &= 0 \qquad \text{(all } \beta\text{)}, \end{aligned} \tag{6.2}$$

where $K_\beta = K_\beta(r)$ is defined by (5.2) and β will be chosen later.

The proof can now be briefly outlined as follows. (6.2) is split into two parts I and Q. I consists of those terms such that

$$q^{k_0} + \cdots + q^{k_r} \leqq q^{\beta} + \cdots + q^{\beta - r},$$

and Q those terms such that

$$q^{k_0} + \cdots + q^{k_r} > q^{\beta} + \cdots + q^{\beta - r}.$$

Then (6.2) gives

$$I + Q = 0. \tag{6.3}$$

We prove *that every term of I is a polynomial* and, for a suitable choice of β, *every term of Q is of negative degree.* This with (6.3) implies

$$I = 0, \qquad Q = 0. \tag{6.4}$$

We proceed to show that $I \neq 0$. We write $I = I_1 + I_2$ where I_1 is the sum of the terms of (6.2) such that

$$q^{k_0} + \cdots + q^{k_r} < q^{\beta}$$

and I_2 those such that

$$q^{\beta} \leqq q^{k_0} + \cdots + q^{k_r} \leqq q^{\beta} + \cdots + q^{\beta - r}.$$

Then, the first equation of (6.4) is

$$I_1 + I_2 = 0. \tag{6.5}$$

Next, it is noted that M, defined by (5.6), divides every term of I_1. On the other hand, it is found that for β sufficiently large, deg M is greater than the degree of I_2. (6.5), therefore, implies

$$I_1 = 0, \qquad I_2 = 0. \tag{6.6}$$

But, for a suitable choice of β, there exists a single non-zero term H of I_2 which has the property that the irreducible polynomial E introduced above divides every other term of I_2 a greater number of times than it divides H. This clearly implies

$$I_2 \neq 0, \tag{6.7}$$

and we have a contradiction of (6.6).

We proceed with the proof of the various parts, outlined above, in order.

(i) *Every term of I is a polynomial.* This follows immediately from Lemmas 5.2 and 5.6 and (5.8).

(ii) *Every term of Q is of negative degree.* Let N be any term of Q and consider its degree.

Case 1. $k_0 \geqq \beta + 1$. From Lemma 5.6

$$\begin{aligned}
\deg N &\leqq \beta q^\beta + \cdots + (\beta - r)q^{\beta-r} + e(q^{k_0} + \cdots + q^{k_r}) \\
&\qquad\qquad + cq^{k_0} - (k_0 q^{k_0} + \cdots + k_r q^{k_r}) \\
&= \beta q^\beta + \cdots + (\beta - r)q^{\beta-r} - (k_0 - e)q^{k_0} \\
&\qquad\qquad - \cdots - (k_r - e)q^{k_r} + cq^{k_0} \\
&\leqq (\beta(q^r + \cdots + 1) - \beta q^{r+1} + \text{const.})q^{\beta-r} \\
&\to -\infty, \quad \beta \to \infty,
\end{aligned} \tag{6.8}$$

where const. denotes a constant independent of β. Therefore, there exists a β_1 such that for $\beta > \beta_1$, the degree of all such terms is less than 0.

Case 2. $k_0 \leqq \beta$. Choose the least i such that

$$q^{k_0} + \cdots + q^{k_i} > q^\beta + \cdots + q^{\beta-r} \qquad (k_0 \geqq \cdots \geqq k_i).$$

Then $q^{k_0} + \cdots + q^{k_{i-1}} \leqq q^\beta + \cdots + q^{\beta-r}$ and by Lemma 5.9 $q^{k_0} + \cdots + q^{k_{i-1}} \leqq q^\beta + \cdots + q^{\beta-i+1} \leqq q^\beta + \cdots + q^{\beta-r+1}$. Therefore, $k_i > \beta - r$ and we may write $k_j = \beta - l_j$ $(j = 0, \cdots, i)$, where $l_i < r$. Hence,

$$\begin{aligned}
\deg N &\leqq \beta q^\beta + \cdots + (\beta - r)q^{\beta-r} + e(q^{k_0} + \cdots + q^{k_r}) + cq^\beta \\
&\qquad\qquad - ((\beta - l_0)q^{k_0} + \cdots + (\beta - l_i)q^{k_i}) \\
&< \beta\{(q^\beta + \cdots + q^{\beta-r}) - (q^{k_0} + \cdots + q^{k_i})\} + l_0 q^{k_0} \\
&\qquad\qquad + \cdots + l_i q^{k_i} + cq^\beta + e(q^{k_0} + \cdots + q^{k_r}) \\
&< -\beta q^{\beta-r} + l_0 q^{k_0} + \cdots + l_i q^{k_i} + cq^\beta + e(q^{k_0} + \cdots + q^{k_r}) \\
&\to -\infty, \quad \beta \to \infty.
\end{aligned} \tag{6.9}$$

Hence, there exists a β_2 such that for $\beta > \beta_2$, $\deg N < 0$.

(iii) *Every term of I_1 is divisible by M.* This follows immediately from Lemma 5.5.

(iv) *Degree of M is greater than degree of I_2.* Let N be any term of I_2. By Lemma 5.10 there exists an i such that $q^{k_0} + \cdots + q^{k_i} = q^\beta$. Then we write $k_j = \beta - l_j$ $(j = 0, \cdots, i)$ and l_j will be $\leqq d$ by Lemma 5.7.

$$\begin{aligned}
\deg M - \deg N &\geqq -(q - 1)((\beta - 1)q^{\beta-1} + \cdots + (\beta - d)q^{\beta-d}) \\
&\qquad + (\beta - l_0)q^{k_0} + \cdots + (\beta - l_i)q^{k_i} - e(q^{k_0} + \cdots + q^{k_r}) - cq^{k_0} \\
&> -(q - 1)\beta(q^{\beta-1} + \cdots + q^{\beta-d}) + \beta q^\beta - dq^\beta - e(r + 1)q^\beta - cq^\beta \\
&\to \infty, \quad \beta \to \infty.
\end{aligned} \tag{6.10}$$

Hence, there exists a β_3 such that for all $\beta > \beta_3$, $\deg M$ is greater than $\deg I_2$.

(v) $I_2 \neq 0$. Write (since $r + 1 = q^s$)

$$H = \frac{K_\beta S(\beta - s, \cdots, \beta - s)}{F_{\beta-s}^{q^s}}$$

$$= \frac{K_\beta(-1)^{\beta-s} E^{q^\beta} C_{r+1}^{q^{\beta-s}}}{F_{\beta-s}^{q^s}} \neq 0.$$

Now choose β so that

$$\beta \equiv s \pmod{e}. \tag{6.11}$$

Let κ be the number of times that E divides K_β. Let h equal the number of times that E divides H. Thus, from Lemma 5.8 and the fact that we chose $e > \deg C_{r+1}$,

$$h = \kappa + q^\beta - q^s(1 + q^e + \cdots + q^{\beta-s-e}). \tag{6.12}$$

If G is any other non-zero term of I_2 and g is the power of E dividing G, then we want to prove that $g > h$.

Case 1. $q^{k_0} + \cdots + q^{k_r} = q^\beta$ *for* G.

K_β contributes κ as before in H. The symmetric function from (6.1) contributes at least q^β as before. Now examine the denominator of G. By (6.11) and Lemma 5.8, the fact that E divides $[l]$ in G implies that $[l]$ is of the form $\beta - s - ve$. By Lemma 5.7, $k_r \geqq \beta - d$. Since we chose $e > d$, we have for $v \geqq 1$ that $\beta - s - ve < \beta - d$; and, therefore, the number of E's dividing the brackets $[\beta - s - ve]$ $(v \geqq 1)$ is the power of $[\beta - s - ve]$ occurring, i.e., $q^{\beta-(\beta-s-ve)} = q^{s+ve}$—the same as for H. There remains only the exponent of $[\beta - s]$ to examine since E cannot divide $[\beta - s + 1], \cdots, [\beta - 1]$ or $[\beta]$. Now $k_r < \beta - s$ or we would otherwise have $q^{k_0} + \cdots + q^{k_r} > q^\beta$. Therefore, the exponent of $[\beta - s] \leqq (q^{k_0} + \cdots + q^{k_{r-1}})q^{-(\beta-s)} < q^s$. Thus

$$g > \kappa + q^\beta - q^s - (q^{s+e} + \cdots + q^{\beta-e}) = h.$$

Case 2. $q^{k_0} + \cdots + q^{k_r} > q^\beta$ *for* G.

By Lemma 5.10 there exists an i such that $q^{k_0} + \cdots + q^{k_i} = q^\beta$. By the argument above these k's contribute at least as many E's as occur in H. Let k_j be any other k. From the symmetric sum we have q^{k_j} E's, while from F_{k_j} in the denominator at most

$$q^{k_j-e} + q^{k_j-2e} + \cdots + 1 < q^{k_j}.$$

Again $g > h$.

Therefore, E^{h+1} does not divide H but divides all other terms of I_2 and from this $I_2 \neq 0$.

Hence for any $\beta > \max(\beta_1, \beta_2, \beta_3)$ and satisfying condition (6.11) we have a contradiction and the proof is complete.

7. $\psi(\alpha)$ **for algebraic** $\alpha \neq 0$. We first need an additional lemma.

LEMMA 7.1. *If $\vartheta = \psi(\alpha)$ is a root of the equation*

$$\sum_{j=l}^{m} A_j t^{q^j} = 0 \qquad (A_l \neq 0,\ A_i \text{ integral}), \tag{7.1}$$

then $\eta = \psi(E\alpha)$, where E is irreducible and deg $E = e >$ max deg A_i, *is a root of an equation*

$$\sum_{j=l}^{m} D_j t^{q^j} = 0 \qquad (D_j \text{ integral}) \tag{7.2}$$

such that

$$D_j \equiv A_j \pmod{E}. \tag{7.3}$$

In particular, $D_l \not\equiv 0 \pmod{E}$.

Proof. From (7.1), we may write

$$A_l \vartheta^{q^l} = -\sum_{j=l+1}^{m} A_j \vartheta^{q^j}. \tag{7.4}$$

By the multiplication theorem (1.4), and since E divides $\psi_j(E)/F_j$ $(j < e)$,[9]

$$\begin{aligned} \eta &= \sum_{j=0}^{e} \frac{(-1)^j}{F_j} \psi_j(E) \vartheta^{q^j} \\ &\equiv (-1)^e \vartheta^{q^e} \qquad \pmod{E}. \end{aligned} \tag{7.5}$$

Then for f sufficiently large, we may write, using (7.4) repeatedly,

$$\begin{aligned} A_l^{q^{fe}} \eta^{q^l} &= (-1)^e A_l^{q^{fe}} \vartheta^{q^{e+l}} + A_l^{q^{fe}} \sum_{j=0}^{e-1} \frac{(-1)^j}{F_j^{q^l}} \psi_j^{q^l}(E) \vartheta^{q^{j+l}} \\ &= \sum_{j=l+1}^{m} B_j^{(l)} \vartheta^{q^{e+j}}, \end{aligned}$$

where the $B_j^{(l)}$ $(j = l + 1, \cdots, m)$ are integral and

$$B_j^{(l)} \equiv (-1)^{e+1} A_j \pmod{E} \qquad (j = l+1, \cdots, m). \tag{7.6}$$

Similarly, we may write

$$\begin{aligned} A_l^{q^{fe-1}} \eta^{q^{l+i}} &= (-1)^e A_l^{q^{fe-1}} \vartheta^{q^{e+l+i}} + A_l^{q^{fe-1}} \sum_{j=0}^{e-1} \frac{(-1)^j}{F_j^{q^{l+i}}} \psi_j^{q^{l+i}}(E) \vartheta^{q^{j+l+i}} \qquad (i = 1, \cdots, m-l) \\ &= \sum_{j=l+1}^{m} B_j^{(l+i)} \vartheta^{q^{e+j}}, \end{aligned}$$

[9] Also $\psi_e(E) = F_e$. By the congruence in (7.5), we mean that the coefficients of like powers of ϑ are congruent (mod E).

where the $B_j^{(l+i)}$ are integral and

$$(7.7) \qquad B_j^{(l+i)} \equiv \begin{cases} 0 & \text{if } j \neq l+i \\ (-1)^e & \text{if } j = l+i \end{cases} \qquad (\text{mod } E).$$

We then seek a set of J's such that

$$(7.8) \qquad A_l^{q^{l}e} J_l \eta^{q^l} + A_l^{q^{l}e-1} J_{l+1} \eta^{q^{l+1}} + \cdots + A_l^{q^{l}e-1} J_m \eta^{q^m} = 0,$$

i.e., we seek a suitable solution of the homogeneous system

$$(7.9) \qquad \sum_{i=l}^{m} J_i B_j^{(i)} = 0 \qquad (j = l+1, \cdots, m).$$

The matrix of (7.9), where each element is taken mod E, is

$$(7.10) \qquad \begin{pmatrix} (-1)^{e+1} A_{l+1} & (-1)^e & 0 & 0 & \cdots & 0 \\ (-1)^{e+1} A_{l+2} & 0 & (-1)^e & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ (-1)^{e+1} A_m & 0 & 0 & 0 & \cdots & (-1)^e \end{pmatrix}.$$

The rank of the matrix (7.10) is $m - l$, and that of (7.9) is then clearly also of rank $m - l$. Hence, J_{l-1+j} can be taken[10] as $(-1)^j$ times the determinant obtained by omitting the j-th column from the matrix of (7.9). In particular, it is seen that

$$J_l \equiv (-1)^{e(m-l)+1} \quad (\text{mod } E),$$

$$J_{l+i} \equiv (-1)^{e(m-l)+1} A_{l+i} \quad (\text{mod } E) \quad (i = 1, \cdots, m-l).$$

The lemma now follows immediately from (7.8).

Now consider $\psi(\mu)$, where $\mu \neq 0$ is algebraic. Suppose that $\psi(\mu)$ is algebraic. There exists a polynomial G such that $G\mu = \alpha$ is an algebraic integer. By the multiplication formula, $\psi(\alpha)$ is also algebraic. By Lemma 2.1, we may suppose that it is a root of (2.1). Let $\alpha = \alpha_0$ be a root of (5.8). Suppose that $r + 1 = q^s$. Let E be a polynomial of deg e that satisfies the following conditions.

(i) E is irreducible.

(ii) $e > \max(\deg C_{r+1}, d, \deg A_i, s + l)$ $(i = l, \cdots, m)$, where d is defined by (5.5).

Place

$$\gamma_i = E\alpha_i \qquad (i = 0, \cdots, r).$$

By Lemma 7.1 $\psi(\gamma) = \psi(E\alpha)$ is a root of the linear polynomial (7.2). We have

$$(7.11) \qquad \sum_{j=l}^{m} D_j \sum_{k=0}^{\infty} \frac{(-1)^k \gamma^{q^{k+j}}}{F_k^{q^j}} = \sum_{k=l}^{\infty} \frac{B_k \gamma^{q^k}}{F_k},$$

[10] See for example M. Bôcher, *Introduction to Higher Algebra*, p. 47.

where

$$B_k = \sum_{j=l}^{m} \frac{(-1)^{k-j} D_j F_k}{F_{k-j}^{q^j}}. \tag{7.12}$$

Therefore,

$$\begin{aligned} K_\beta \prod_\gamma \sum_{k=l}^{\infty} \frac{B_k \gamma^{q^k}}{F_k} &= K_\beta \sum_{k_0 \geqq \cdots \geqq k_r} \frac{B_{k_0} \cdots B_{k_r}}{F_{k_0} \cdots F_{k_r}} S(k_0, \cdots, k_r) \\ &= 0 \qquad \text{(all } \beta), \end{aligned} \tag{7.13}$$

where

$$\begin{aligned} S(k_0, \cdots, k_r) &= \sum \gamma_0^{q^{k_0}} \cdots \gamma_r^{q^{k_r}} \\ &= E^{q^{k_0} + \cdots + q^{k_r}} \sum \alpha_0^{q^{k_0}} \cdots \alpha_r^{q^{k_r}}, \end{aligned} \tag{7.14}$$

the sums being symmetric sums of the indicated letters.

Now define I, I_1, I_2, M, and Q exactly as in the proof for ξ. The outline of the proof is exactly the same; the details, however, are more complex.

(i) *Every term of I is a polynomial.* This follows immediately from Lemmas 5.2 and 5.6 and (5.8), if we note from (7.12) that B_k is always integral.

(ii) *Every term of Q is of negative degree.*

(iii) *The degree of M is greater than the degree of I_2.* From (7.12) we note that for β sufficiently large

$$\deg B_\beta = mq^\beta + \deg D_m. \tag{7.15}$$

Also (7.13) differs from (6.2) only by the $r + 1$ B's. It is then easy to verify that we have the same limits as in (6.8), (6.9), (6.10).

(iv) *Every term of I_1 is divisble by M.* This follows immediately from Lemma 5.5.

(v) $I_2 \neq 0$. Place

$$\begin{aligned} H &= \frac{K_\beta B_{\beta-s}^{q^s} S(\beta - s, \cdots, \beta - s)}{F_{\beta-s}^{q^s}} \\ &= \frac{K_\beta B_{\beta-s}^{q^s} C_{r+1}^{q^{\beta-s}}}{F_{\beta-s}^{q^s}} \neq 0. \end{aligned} \tag{7.16}$$

We show that E divides every other term of I_2 more times than it divides H. Now choose β so that

$$\beta \equiv s + l \pmod{e}. \tag{7.17}$$

Therefore, from (7.12), Lemma 7.1 and Lemma 5.8,

$$B_{\beta-s} \equiv \frac{(-1)^{\beta-s-l} D_l F_{\beta-s}}{F_{\beta-s-l}^{q^l}} \not\equiv 0 \pmod{E}. \tag{7.18}$$

Let κ be the number of times E divides K_β and h the number of times E divides H. Thus, from Lemma 5.8, (7.17) and (7.18),

$$h = \kappa + q^\beta - q^e(q^l + q^{l+e} + \cdots + q^{\beta-s-e}).$$

If G is any other non-zero term of I_2 and g is the power of E that divides G, then we want to prove $g > h$.

Case 1. $q^{k_0} + \cdots + q^{k_r} = q^\beta$ *for* G.

K_β contributes κ as in H. The symmetric function contributes at least q^β E's, the number it contributed to h. Since $e > s + l$, $k_0 < \beta - s - l + e$. Hence $F_{k_0} \cdots F_{k_r}$ contains at most

$$(q^{k_0} + \cdots + q^{k_r})q^{-(\beta-s-l)} + \cdots + (q^{k_0} + \cdots + q^{k_r})q^{-e} = q^{s+l} + \cdots + q^{\beta-e}$$

E's, the same as for h.

Either

$$q^{\beta-s} > q^{k_r} \geqq q^{\beta-s-l} \tag{7.19}$$

or

$$q^{k_r} < q^{\beta-s-l}, \tag{7.20}$$

since $k_r \geqq q^{\beta-s}$ implies that $G = H$ or $q^{k_0} + \cdots + q^{k_r} > q^\beta$. If (7.19) holds,

$$B_{k_r} \equiv 0 \pmod{E}$$

because B_{k_r} by (7.12) is divisible by

$$[k_r][k_r - 1]^q \cdots [k_r - l + 1]^{q^{l-1}}.$$

Then $g > h$.

If (7.20) holds, $F_{k_0} \cdots F_{k_r}$ contains at most

$$(q^{k_0} + \cdots + q^{k_{r-1}})q^{-(\beta-s-l)} + (q^{k_0} + \cdots + q^{k_r})(q^{-(\beta-s-l-1)} + \cdots + q^{-e})$$
$$= (q^\beta - q^{k_r})q^{-(\beta-s-l)} + q^{s+l+1} + \cdots + q^{\beta-e} < q^e(q^l + \cdots + q^e),$$

and $g > h$.

Case 2. $q^{k_0} + \cdots + q^{k_r} > q^\beta$ *for* G.

By Lemma 5.10 there exists an i such that $q^{k_0} + \cdots + q^{k_i} = q^\beta$. By the argument above these k's contribute at least as many E's as occur in H. Let k_j be any other k. From the symmetric sum we have q^{k_j} E's, while from F_{k_j} in the denominator at most

$$q^{k_j-e} + q^{k_j-2e} + \cdots + 1 < q^{k_j}.$$

Again $g > h$.

Therefore, in every case E^{h+1} does not divide H but divides all other terms of I_2 and from this $I_2 \neq 0$.

Therefore, for any β sufficiently large that satisfies (7.17) we arrive at a contradiction. We have

THEOREM 7.1. *$\psi(\alpha)$ is transcendental for $\alpha \neq 0$ algebraic.*

This gives the

COROLLARY. *$\lambda(\alpha)$ is transcendental for $\alpha \neq 0$ algebraic.*

Proof. $\psi(\lambda(\alpha)) = \alpha$.

COROLLARY. *If G and H are polynomials and $\alpha \neq 0$ is algebraic, $G\lambda(\alpha) + H\xi$ is transcendental if not both G and H are zero.*

The question naturally arises as to the possibility of a direct proof of the transcendence of

$$\lambda(1) = \sum_{k=0}^{\infty} \frac{1}{L_k} \tag{7.21}$$

using the series form (7.21). This can be done and suggests consideration of series of the form

$$\sum_{k=0}^{\infty} \frac{B_k}{L_k^{\gamma}} \qquad (\gamma > 0).$$

These and other questions will be left for another paper.

DUKE UNIVERSITY.

THE SUM OF THE DIVISORS OF A POLYNOMIAL

By E. F. Canaday

1. Introduction. Let

$$A = A(x) = x^k + \alpha x^{k-1} + \cdots + \lambda$$

denote a polynomial with coefficients modulo 2. All coefficients may then be written 1 or 0 and the number of polynomials of degree k is 2^k.

Let $\sigma(A)$ denote the sum of the divisors of A. If $\sigma(A_1) = A_2$, we write $A_1 \rightarrow A_2$. Clearly A_1 and A_2 are of equal degree. Consider the sequence $A_1 \rightarrow A_2 \rightarrow A_3 \rightarrow \cdots$ where $\sigma(A_j) = A_{j+1}$. Since the number of polynomials of any given degree is finite, after a certain point an A_i in the sequence will be repeated. If $A_1 \rightarrow A_2 \rightarrow \cdots \rightarrow A_n \rightarrow A_1$, all A_i being distinct, the set $A_1, \cdots, A_n$ will be called an n-ring. In particular, if $A \rightarrow A$, i.e. $\sigma(A) = A$, we shall call A a one-ring. Other definitions and notation are given in section 2.

Section 3 contains theorems on weight. In section 4 are developed some invariant properties of the operator σ. Section 5 consists principally of lemmas needed for subsequent theorems. It also contains the useful theorem: *The only complete polynomials whose irreducible factors are all of the form* $x^\alpha(x + 1)^\beta + 1$ *are* $x^2 + x + 1$, $x^4 + x^3 + x^2 + x + 1$ *and* $x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$.

In sections 6–8, one-rings and methods of constructing them are discussed. First the trivial type $x^{2^n-1}(x + 1)^{2^n-1}$ is treated, then the type $x^h(x + 1)^k A$, of which eleven are found, and then a proof is given that there are no others of certain sub-types. Lastly the type B^2, where $(B, x(x + 1)) = 1$, is discussed but none found. It seems plausible that none of this type exist but this is not proved.

Section 9 is devoted mainly to two-rings. There is the infinite class

$$x^{2^\alpha-1}(x + 1)^{2^\beta-1} \leftrightarrow x^{2^\beta-1}(x + 1)^{2^\alpha-1} \qquad (\alpha \neq \beta)$$

corresponding to the infinite class of trivial one-rings. In addition we determine the two-rings of certain forms. The simplest of these is

$$x^\alpha(x + 1)^\beta \rightarrow A \rightarrow x^\alpha(x + 1)^\beta. \tag{1.1}$$

We show that there are only three of these.

Generalizing (1.1) we seek all rings of the form

$$x^{\alpha_1}(x + 1)^{\beta_1} \rightarrow A_1 \rightarrow x^{\alpha_2}(x + 1)^{\beta_2} \rightarrow A_2 \rightarrow \cdots \rightarrow A_p \rightarrow x^{\alpha_1}(x + 1)^{\beta_1}, \tag{1.2}$$

where alternate polynomials are of the form $x^\alpha(x + 1)^\beta$. We show that there are only nine of this form, three two-rings (1.1) and six four-rings.

Received September 20, 1941.

Tables have been prepared and are available at the Duke University library showing the polynomials, 2046 in number, of degree $\leqq 10$, in both factored and expanded form, also the weight and the sum of the factors of each. For each value of k there is shown also the number of rings, number of polynomials in each ring and branch (defined in section 2) and the number of polynomials of each weight in each branch.

2. Definitions and notation. The letters $A, B, C, \cdots$ will be used to represent arbitrary polynomials mod 2; the letters $P, Q, R, \cdots$ will represent irreducible polynomials—usually of degree $\geqq 2$. The degree of A will be denoted by deg A.

Corresponding to formulas for ordinary integers, we have

$$\sigma(P^n) = P^n + P^{n-1} + \cdots + 1,$$

$$\sigma(AB) = \sigma(A)\sigma(B) \qquad \text{for } (A, B) = 1,$$

where (A, B) denotes the "greatest" common divisor of A and B.

An n-ring, defined in section 1, will be symbolized by $\mathfrak{R}^n$. In particular, if $A \to B$ and $A = B$, this will be written $A \rightrightarrows B$. If $A \to B$ and $B \to A$, where $A \neq B$ we have an $\mathfrak{R}^2$ and this will be written $A \leftrightarrow B$.

A ring together with all polynomials which lead into it will be called a *branch*.

If $A_i \to B$ $(i = 1, \cdots, n)$ but no other $A \to B$, then B is said to be of *weight n*. If no $A \to B$, then B is of zero weight.

The polynomial $x^k + x^{k-1} + \cdots + x + 1$ will be called *complete*. More generally, $A^k + A^{k-1} + \cdots + A + 1$ is complete in A.

If $A = A(x)$ is of degree m, and $A^* = A^*(x) = x^m A\left(\frac{1}{x}\right)$, then we say A *inverts* into A^*, and consequently A^* inverts into A. Thus any complete polynomial inverts into itself.

3. Theorems on weight. Put

$$A = x^\alpha(x + 1)^\beta \Pi P_i^{\alpha_i} \qquad (i = 1, \cdots, n), \tag{3.1}$$

$$A \to (x^\alpha + x^{\alpha-1} + \cdots + 1)[(x + 1)^\beta + \cdots + 1] \Pi(P_i^{\alpha_i} + \cdots + P_i + 1) = A_1 \cdot \tag{3.2}$$

If k is even and A is not a perfect square, at least one exponent of A is odd. If α_i is odd the corresponding factor of A_1 is divisible by $x(x + 1)$. If every α_i is even, then α and β are odd and $x(x + 1)$ divides the product of the first two factors of A_1.

Now consider A, a perfect square. Since k is even (3.1) becomes

$$A = x^{2\alpha}(x + 1)^{2\beta} \Pi P_i^{2\alpha_i} \qquad (i = 1, 2, \cdots, n),$$

and (3.2) becomes

$$A \to (x^{2\alpha} + \cdots + 1)[(x + 1)^{2\beta} + \cdots + 1]\Pi(P_i^{2\alpha_i} + \cdots + P_i + 1) = A_1 .$$

Here $x \nmid A$ and $x + 1 \nmid A$, hence

THEOREM 1. *For k even all polynomials of weight greater than zero are divisible by both or neither of the terms x and $x + 1$. Every such polynomial divisible by neither factor is the sum of the divisors of a perfect square.*

As a consequence of this theorem a sufficient condition that a polynomial A of even degree be of zero weight is that A be divisible by only one of the factors x and $x + 1$, that is, $A(0) + A(1) = 1$.

Next consider k odd and $A_1 = xB_1^2$, $A_2 = (x + 1)B_2^2$.

$$A_1 \to A_1', \qquad x + 1 \mid A_1' \text{ but } x \nmid A_1',$$

$$A_2 \to A_2', \qquad x \mid A_2' \qquad \text{but } x + 1 \nmid A_2'.$$

Now with k still odd, suppose $A_1 \neq xB_1^2$ or $(x + 1)B_2^2$, that is, $A_1 = x^{\alpha}(x + 1)^{\beta}\Pi P_i^{\alpha_i}$, where α and β are both odd or some α_i is odd. If α and β are both odd, then some α_i is odd and $x(x + 1) \mid P_i$. We may now state

THEOREM 2. *If k is odd, all A of weight greater than zero are divisible by $x(x + 1)$ except those which are the sums of factors of polynomials of the form $x \cdot B^2$ or $(x + 1)B^2$, and these are divisible by x or $x + 1$ but not by both.*

As a consequence we have

COROLLARY 1. *All irreducible polynomials of odd degree are of zero weight.*

Polynomials of a given degree may be divided into four equal sets as follows: (1) those divisible by x but not by $(x + 1)$; (2) those divisible by $(x + 1)$ but not by x; (3) those divisible by both x and $(x + 1)$; (4) those divisible by neither.

For k even, sets (1) and (2) have been shown to be of zero weight. Set (4) is zero weight except for those polynomials, α in number, which are the sums of the factors of perfect squares. Polynomials in set (3) are of positive weight except for a certain set of zero weight. A general method for determining the exact number, N, of this set has not been found. Thus the number of polynomials of even degree k, and of zero weight is

$$2^{k-1} + 2^{k-2} - \alpha + N.$$

For $k = 2, 4, 6$, $N = 0$. For $k = 8$, $N = 9$.

For k odd, similar reasoning and Theorem 2 show sets (1) and (2) are zero weight except for those polynomials which are sums of factors of polynomials, β in number, of form $x \cdot B^2$ and $(x + 1)B^2$, where B contains no first degree factor other than the one by which it is multiplied. Set (4) is zero weight. Set (3) is positive weight except for a certain set M which are of zero weight. Thus the total number of zero weight is

$$2^{k-2} + 2^{k-1} - \beta + M.$$

THEOREM 3. *Only polynomials which are powers of x, and perfect squares, have as the sums of their factors polynomials not divisible by x.*

This follows easily from $\sigma x^n = x^n + \cdots + x + 1$ which is not divisible by x and $\sigma P^{2k+1} = P^{2k+1} + \cdots + P + 1$ which is divisible by $P + 1$ and hence by x (here P denotes $x + 1$ or an irreducible of degree $\geqq 2$).

Now consider $A_1, A_2, A_3, \cdots, A_w \to A$ and no other $A_i \to A$. Then A has weight w.

Let B be prime to all the A's, $A_nB \to A\sigma B$ $(n = 1, 2, 3, \cdots, w)$ and the weight of $A\sigma B \geqq w$. Hence we may state

THEOREM 4. *If w is the weight of any polynomial of degree k and N is the number of irreducibles of degree $k_1 > k$, then there are at least N polynomials of degree $k + k_1$ and weight $\geqq w$.*

THEOREM 5. *Given polynomial A of degree $2k$ and $A(0) + A(1) = 1$ so that A is of zero weight. Then A^n is also of zero weight.*

Proof. $A(0) = A^n(0)$ and $A(1) = A^n(1)$. Therefore $A(0) + A(1) = A^n(0) + A^n(1) = 1$ and hence A^n is of zero weight.

4. Invariant property of operator σ. We now develop an invariant property of the operator σ when applied to the polynomial $A = A(x) = x^k + \alpha_1 x^{k-1} + \cdots + \alpha_k$. We define $\alpha(A) = \alpha_1$ and $f(A) = \alpha_2 + \alpha_3 + \cdots + \alpha_{k-1}$; then $\alpha(A) + f(A) = A(0) + A(1) + 1$.

(1) If $x \nmid A$ and $x + 1 \nmid A$, then $\alpha(\sigma A) = \alpha A$.

(2) If $A = x^m(x + 1)^n B$, $(B, x(x + 1)) = 1$, $m \geqq 1$, $n \geqq 1$,

$$\alpha(\sigma A) = 1 + n + 1 + \alpha(\sigma B) = \alpha(B) + n,$$

$$\alpha(A) = \alpha[x^m(x^n + nx^{n-1} + \cdots + 1)B] = \alpha(B) + n,$$

hence $\alpha(\sigma A) = \alpha(A)$.

(3) If $A = x^n B$, $(B, x(x + 1)) = 1$, $n \geqq 1$, then $\alpha(A) = \alpha(B)$ and $\sigma A = (x^n + \cdots + x + 1)\sigma B$,

$$\alpha(\sigma A) = \alpha(\sigma B) + 1 = \alpha(B) + 1 = \alpha(A) + 1.$$

(4) If $A = (x + 1)^n B$, $(B, x(x + 1)) = 1$, $n \geqq 1$, then

$$\alpha(A) = \alpha(B) + n,$$

$$\sigma A = [x^n + (n + 1)x^{n-1} + \cdots + 1]\sigma B,$$

$$\alpha(\sigma A) = \alpha(\sigma B) + n + 1 = \alpha(A) + 1.$$

Polynomials will be said to belong to class I if $f(A) = 1$, and to class II if $f(A) = 0$. Thus in (1) and (2) above, A and σA belong to the same class; in cases (3) and (4) they belong to different classes.

(5) If $A = \Pi P_i^{l_i}$, $(A, x(x + 1)) = 1$, then

$$(\sigma A)_0 = \Pi(P_i^{l_i} + \cdots + P_i + 1)_0 = 1$$

if all l_i are even, and $= 0$ otherwise. Likewise $(\sigma A)_1 = 0$ if at least one l_i is odd and $= 1$ if all l_i are even.

(6) If $A = x^{\epsilon}\Pi P_i^{l_i}$, we have

$$(\sigma A)_0 = \Pi(P_i^{l_i} + \cdots + P_i + 1)_0 = \Pi(l_i + 1) = 0,$$

unless all l_i are even. That is, if $x + 1 \nmid A$, then $(\sigma A)_1 = \Pi l_i = 0$ unless all l_i are even.

(7) If $A = (x + 1)^{\eta}\Pi P_i^{l_i}$, then $(\sigma A)_1 = 0$ unless all l_i are even.

(8) If $A = x^{\epsilon}(x + 1)^{\eta}\Pi P_i^{l_i} = x^{\epsilon}(x + 1)^{\eta}B$, then

$$(\sigma A)_0 = (\eta + 1)\Pi(l_i + 1) \quad \text{and} \quad (\sigma A)_1 = (\epsilon + 1)\Pi(l_i + 1),$$

$$(\sigma A)_0 + (\sigma A)_1 = (\epsilon + \eta)\Pi(l_i + 1).$$

If $\epsilon = \eta = 0$, then $(\sigma A)_0 + (\sigma A)_1 = 0$ and

$$\alpha(\sigma A) + f(\sigma A) = 1,$$

$$\alpha(A) + f(A) = 1 + A(0) + A(1) = 1,$$

$$f(\sigma A) = f(A).$$

If $\epsilon > 0$ and $\eta = 0$, then $A(0) = 0$, $A(1) = 1$, $\alpha(A) + f(A) = 0$,

$$\alpha(\sigma A) + f(\sigma A) = 1 + \epsilon\Pi(l_i + 1),$$

$$f(\sigma A) = f(A) + \epsilon\Pi(l_i + 1).$$

Likewise if $\epsilon = 0$ and $\eta > 0$, then $A(0) = 1$ and $A(1) = 0$; hence

$$f(\sigma A) = f(A) + \eta\Pi(l_i + 1).$$

If $\epsilon > 0$ and $\eta > 0$, then $\alpha(A) + f(A) = 1$,

$$\alpha(\sigma A) + f(\sigma A) = 1 + (\epsilon + \eta)\Pi(l_i + 1),$$

$$f(\sigma A) = f(A) + (\epsilon + \eta)\Pi(l_i + 1).$$

We may therefore state the following

THEOREM 6. *For polynomials A of even degree, $f(\sigma A) = f(A)$; that is, the divisor sum belongs to the same class as the polynomial itself.*

We also have

THEOREM 7. *For polynomials of arbitrary degree,*

$$A = x^{\epsilon}(x + 1)^{\eta}\Pi P_i^{l_i},$$

(a) *if $(\epsilon + \eta)$ is even, then $f(\sigma A) = f(A)$,*

(b) *if at least one l_i is even, $f(\sigma A) = f(A)$;*

hence if either (a) *or* (b) *is true, A and σA are polynomials of the same class.*

Relations between $\beta(A)$ and $\beta(\sigma A)$, where $\beta(A)$ represents the coefficient of x^{k-2}, have been obtained but for the sake of brevity are omitted.

5. Lemmas.

LEMMA 1. *If* $A = x^{h-1} + x^{h-2} + \cdots + 1$ *is a complete polynomial and* $(x + 1)^r$ *divides A but* $(x + 1)^{r+1}$ *does not, then* $r = 2^n - 1$ *and* $A = (x + 1)^{2^n-1}B^{2^n}$ *where B is complete.*

LEMMA 2. $P = x(x + 1)^{2^m-1} + 1$ *is irreducible only for* $m = 1$ *and* $m = 2$.

Proof. Since $P = \dfrac{x^{2^m+1} + 1}{x + 1}$, we have $x^{2^m+1} \equiv 1 \pmod P$. From this it follows that $x^{2^{2m}-1} \equiv 1 \pmod P$.

Then the degrees of the irreducible divisors of $x^{2^{2m}} - x$ are divisors of $2m$. (Dickson, *Linear Groups*, 1901, p. 16.) Hence for P to be irreducible it is necessary that $2^m \leqq 2m$, which implies $m \leqq 2$. As a consequence of this lemma we have

COROLLARY. *The only complete irreducibles of the form* $x(x + 1)^\beta + 1$ *are* $x^2 + x + 1$ *and* $x^4 + x^3 + x^2 + x + 1$.

LEMMA 3. *If* $P = x(x + 1)^{2^m} + 1 = x^{2^m+1} + x + 1$ *is irreducible, then* $m \leqq 3$.

Proof. If $P = x^{2^m+1} + x + 1$, then $x^{2^m+1} \equiv x + 1 \pmod P$ and $(x^{2^m+1})^{2^m} \equiv x^{2^m} + 1$. Multiplying by x we have

$$x^{2^{2m}+2^m+1} \equiv x^{2^m+1} + x \equiv 1.$$

This implies $x^{2^{3m}} \equiv x$ and therefore the irreducible factors of P are of degree $\leqq 3m$.

Then for P to be irreducible it is necessary that $3m \geqq 2^m + 1$ and hence $m \leqq 3$.

It follows that the only irreducibles $P = x(x + 1)^{2^m} + 1$ are

$$x(x + 1)^2 + 1 \quad \text{and} \quad x(x + 1)^8 + 1.$$

LEMMA 4. *If* $PQ = x^{2h} + \cdots + 1$ *and* $P = (x + 1)^{2k} + \cdots + 1$, *then* $h = 4$ *and* $k = 1$; *that is,* $P = x^2 + x + 1$, $Q = x^6 + x^3 + 1$, $PQ = x^8 + \cdots + 1$.

Proof. Put $x^{2h} + \cdots + 1 = PQ = P^*Q^*$, where $P^* = P\left(\dfrac{1}{x}\right)x^{2k}$ and $Q^* = Q\left(\dfrac{1}{x}\right)x^{2h-2k}$. Then either

(i) $P = Q^*$ and $P^* = Q$, or
(ii) $P = P^*$ and $Q = Q^*$.

Case (i). $Q = P^*$ and $\deg Q = \deg P = 2k$, $2h = 4k$. Then $P = (x + 1)^{2k} + \cdots + 1$, $Q = (x + 1)^{2k} + x(x + 1)^{2k-1} + \cdots + x^{2k}$, and

$$x^{4k} + \cdots + 1 = \frac{(x + 1)^{2k+1} + 1}{x} \cdot \left[(x + 1)^{2k+1} + x^{2k+1}\right]. \tag{5.1}$$

This is impossible since

$$(x+1)^{2k+1} = (x+1)(x^2+1)^k = (x+1)(x^{2k} + kx^{2k-1} + \cdots + kx^2 + 1)$$
$$= x^{2k+1} + x^{2k} + \cdots + kx^3 + kx^2 + x + 1.$$

Then

$$\frac{(x+1)^{2k+1}+1}{x} = x^{2k} + \cdots + kx^2 + kx + 1.$$

Now the second factor in (5.1) can be written

$$(x+1)^{2k+1} + x^{2k+1} = x^{2k} + \cdots + kx^2 + x + 1.$$

Then the product in (5.1) becomes

$$x^{4k} + \cdots + kx^2 + (k+1)x + 1,$$

which is not complete.

Case (ii). $P = P^*$, $Q = Q^*$.

$$P = (x+1)^{2k} + \cdots + 1, \qquad P^* = (x+1)^{2k+1} + x^{2k+1},$$
$$(x+1)^{2k+2} = x^{2k+2} + 1,$$

and hence $2k + 2 = 2^m$.

From this we get $(x+1)^{2^m} \equiv (x+1) \pmod{P}$. Hence $x^{2^m} \equiv x \pmod{P}$, or $P \mid (x^{2^m} + x)$. Then P can be irreducible only when $\deg P = 2^m - 2 \leqq m$, i.e., $m \leqq 2$. The only admissible value is $m = 2$, so that $P = x^2 + x + 1$. This gives

$$Q = 1 + x^3 + \cdots + x^{6s} = \frac{1 + x^{3(2s+1)}}{1 + x^3}.$$

If Q is irreducible, $2s + 1$ is a prime. Now

$$x^{3(2s+1)} \equiv 1 \pmod{Q}. \tag{5.2}$$

Assume $s > 1$, then by Fermat's Theorem

$$2^{2s} - 1 \equiv 0 \pmod{3(2s+1)} \tag{5.3}$$

and by (5.2) and (5.3), $x^{2^{2s}-1} \equiv 1 \pmod{Q}$. This would imply $\deg Q \leqq 2s$, a contradiction. Therefore Q is factorable if $2s + 1 > 3$. Hence the only possibility is $Q = 1 + x^3 + x^6$ and the proof is complete.

LEMMA 5.

$$P^{2a} + P^{2a-1} + \cdots + P + 1 = Q^m \tag{5.4}$$

is impossible if $m > 1$.

Obviously m cannot be even for then the left member contains odd powers of x, while the right does not. Now (5.4) may be written

$$Q^m = (P^{\alpha-1} + \cdots + 1)(P^{\alpha+1} + P) + 1,$$

so that Q is prime to $P^{\alpha-1} + \cdots + P + 1$. Differentiate (5.4) and we get

$$P'(P^{\alpha-1} + \cdots + P + 1)^2 = Q'Q^{m-1}$$

and therefore $Q^{m-1} \mid P'$.

Now $\deg P \geqq \deg P' \geqq (m - 1) \deg Q$ and

$$2\alpha \deg P > 2\alpha(m - 1) \deg Q.$$

$m \deg Q = 2\alpha \deg P$ from (5.4). Then

$$m \deg Q > 2\alpha(m - 1) \deg Q, \qquad m > 2\alpha(m - 1),$$

which is impossible.

Lemma 6. *If* $P^{2\alpha} + \cdots + P + 1 = Q^m A$, $m > 1$, *then* $\deg P > (m - 1) \deg Q$ *if* m *is odd, and* $\deg P > m \deg Q$ *if* m *is even.*

Proof. Differentiating above equation we get

$$P'(P^{\alpha-1} + \cdots + 1)^2 = Q^{m-1}(mQ'A + QA') \tag{5.5}$$

and this implies $Q^{m-1} \mid P'$ so that $(m - 1) \deg Q < \deg P$. If m is even (5.5) yields $Q^m \mid P$ and $m \deg Q < \deg P$.

Lemma 7. *If an irreducible* $P = x^\alpha(x + 1)^\beta + 1$ *inverts into itself, then* $\alpha = 1$, $\beta = 2^m - 1$ *and* $P = x(x + 1)^{2^m-1} + 1$.

Proof. Since $P = x^\alpha(x + 1)^\beta + 1 = (x + 1)^\beta + x^{\alpha+\beta}$ and $(x^\alpha + 1)(x + 1)^\beta = x^{\alpha+\beta} + 1$, $(x^\alpha + 1) \mid (x^{\alpha+\beta} + 1)$ and so $\alpha \mid (\alpha + \beta)$, $\alpha \mid \beta$. However, $(\alpha, \beta) = 1$; hence $\alpha = 1$. Then $(x + 1)^{\beta+1} = x^{\beta+1} + 1$, so $\beta = 2^m - 1$ and $P = x(x + 1)^{2^m-1} + 1$. Then P is complete and by Lemma 2 we have the

Corollary. *If an irreducible* $P = x^\alpha(x + 1)^\beta + 1$ *inverts into itself, then either* $P = x^2 + x + 1$ *or* $P = x^4 + x^3 + x^2 + x + 1$.

Lemma 8. *If an irreducible* $P = x^\alpha(x + 1)^\beta + 1$ *inverts into* $P^* = x^\gamma(x + 1)^\delta + 1$ *and* $P \neq P^*$, *then* $P = x^{2^m}(x + 1) + 1$ *and* $P^* = x(x + 1)^{2^m} + 1$ *or vice versa.*

Proof.

$$x^\alpha(x + 1)^\beta + 1 = (x + 1)^\delta + x^{\gamma+\delta}, \tag{5.6}$$

$$x^\gamma(x + 1)^\delta + 1 = (x + 1)^\beta + x^{\alpha+\beta}. \tag{5.7}$$

If β is odd, (5.6) implies $\gamma = 1$. Adding (5.6) and (5.7) and simplifying, we have

$$(x^\alpha + 1)(x + 1)^\beta = (x + 1)^{\delta+1} = (x + 1)^{\alpha+\beta}.$$

$$x^\alpha + 1 = (x + 1)^\alpha \quad \text{and} \quad \alpha = 2^m.$$

If $m = 0$, we have $\alpha = 1$ and $\beta = \delta$. Then $P = P^*$ as in Lemma 7. If $\alpha > 1$, since $\delta + 1 = \alpha + \beta$ we have $\delta > \beta$, and (5.6) now implies $(x + 1)^\beta / (x^{\delta+1} + 1)$. But, $\delta + 1 = 2^m + \beta$ is odd and $\beta = 1$, $\delta = \alpha$; hence

$$P = x^{2^m}(x + 1) + 1, \qquad P^* = x(x + 1)^{2^m} + 1.$$

If β is even, α is odd, then writing (5.6) as

$$1 + (x + 1)^\delta + x^{\gamma+\delta} = x^\alpha(x + 1)^\beta,$$

since the smallest positive exponent of x on the right is odd, δ is odd and $\alpha = 1$, and we get P and P^* as above but with their values interchanged.

Now by Lemma 3 we have $m \leqq 3$ so that the pairs of values of P and P^* which are irreducible are

$$P = x^3 + x^2 + 1, \qquad P^* = x^3 + x + 1,$$
$$P = x^9 + x^8 + 1, \qquad P^* = x^9 + x + 1.$$

THEOREM 8. *The only complete polynomials* $A = x^{2m} + \cdots + 1$ *whose irreducible factors are of the form* $x^\alpha(x + 1)^\beta + 1$ *are* $x^2 + x + 1$, $x^4 + x^3 + x^2 + 1$ *and* $x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$.

Proof. By Lemma 6 no P^2 divides A. Since A inverts into itself, each irreducible factor inverts into itself or is paired with a second and each inverts into the other. Therefore by Lemmas 7 and 8, A may be written

$$(5.8) \qquad \begin{aligned} A = (x^2 + x + 1)^\alpha(x^4 + x^3 + x^2 + x + 1)^\beta(x^3 + x + 1)^\gamma \\ \cdot (x^3 + x^2 + 1)^\delta(x^9 + x^8 + 1)^\epsilon(x^9 + x + 1)^\eta, \end{aligned}$$

where the exponents $\alpha, \beta, \gamma, \delta, \epsilon, \eta$ are 0 or 1. Multiplication of the various combinations shows that except for the quadratic and quartic the only value of A is

$$(x^3 + x + 1)(x^3 + x^2 + 1) = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1.$$

COROLLARY. *If any polynomial A inverts into itself, and its irreducible factors are all of the form* $x^\alpha(x + 1)^\beta + 1$, *then A is necessarily of the form* (5.8).

LEMMA 9. *If* $P_1^{2^{n-1}}A$ *is a one-ring and if* $P_1^{2^n} \to P_2$, *where* P_1, P_2 *are prime to* A, *then* $P_1^{2^n}P_2A$ *is also a one-ring.*

Note. Contrary to usual supposition, here P_1 and P_2 may be of first degree.

Proof. We are given

$$(5.9) \qquad P_1^{2^{n-1}}A \Rightarrow (P_1^{2^{n-1}} + \cdots + 1)\sigma A,$$

and

$$(5.10) \qquad P_1^{2^n} + \cdots + 1 = P_2.$$

To prove $P_1^{2^n}P_2A$ a one-ring it is only necessary to prove

$$(5.11) \qquad P_1^{2^n}P_2A = (P_1^{2^n} + \cdots + 1)(P_2 + 1)\sigma A.$$

Multiply (5.9) by P_1P_2 and in the result substitute from (5.10) and we get (5.11). Hence $P_1^{2^n}P_2A$ is a one-ring.

The proof of the following lemmas is direct and will be omitted.

LEMMA 10. *If* $x^h(x+1)^kP^lQ^{2^n-1}$, *where* $l \neq 2^r - 1$, *is a one-ring, then* $2n-1 = 2^r - 1$.

LEMMA 11. *If* $x^h(x+1)^kP^{2l}Q^{2^n-1}$ *is a one-ring, then* $2l = 2^m$ *and* $m = n$.

LEMMA 12. *If* $x^h(x+1)^k\Pi P$ *is a one-ring and if* h (*or* k) *is odd, then* $h = 2^n - 1$ (*or* $k = 2^m - 1$).

LEMMA 13. *If* $x^h(x+1)^k\Pi P$ *is a one-ring and* h (*or* k) *is even, then* σx^h (*or* $\sigma(x+1)^k$ *is not an irreducible polynomial unless* h (*or* k) *equals* 2^r.

LEMMA 14. *It is impossible to have* $\sigma x^{2k} = \sigma P^2$ *or, more generally,*

$$\sigma Q^{2m} = \sigma P^{2n}.$$

6. **Construction of one-rings.** Since

$$x^{2^n-1}(x+1)^{2^n-1} \to (x+1)^{2^n-1}x^{2^n-1} \qquad \text{for } n = 1, 2, 3, \cdots,$$

there are infinitely many one-rings of this trivial type.

All other one-rings must be of one of the two types:

(a) $x^h(x+1)^kA$,

(b) B^2, where $(B, x(x+1)) = 1$.

We first show how one-rings of type (a) may be constructed by applying Lemma 9 to the trivial type.

Lemma 9 states that if AP^{2^n-1} is an $\mathfrak{R}'$ and $P^{2^n} \to Q$, $(Q, A) = 1$, then $AP^{2^n}Q$ is also an $\mathfrak{R}'$. Here both P and Q are irreducible.

First consider $AP^{2^n-1} = x(x+1) = \mathfrak{R}'$, where either factor x or $x+1$ may be A. Since $x^2 \to x^2 + x + 1$, we have the one-ring

$$x^2(x+1)(x^2+x+1) = \mathfrak{R}_1'.$$

Now set $AP^{2^n-1} = \mathfrak{R}_1'$, where $P = x^2 + x + 1$. Since

$$(x^2+x+1)^2 \to x^4 + x + 1,$$

we get

$$x^2(x+1)(x^2+x+1)^2(x^4+x+1) = \mathfrak{R}_3'.$$

Since $(x^4+x+1)^2 \to B$, where B is not irreducible, no new $\mathfrak{R}'$ may be obtained from $\mathfrak{R}_3'$.

Now begin with $x^3(x+1)^3 = \mathfrak{R}'$. Since $x^4 \to x^4 + x^3 + x^2 + x + 1$, we have

$$x^4(x+1)^3(x^4+x^3+x^2+x+1) = \mathfrak{R}_6'.$$

In like manner from $\mathfrak{R}_6'$ we may get

$$x^4(x+1)^4(x^4+x^3+1)(x^4+x^3+x^2+x+1) = \mathfrak{R}_7'.$$

Next begin with

$$x^5(x+1)^3(x^3+x+1)(x^3+x^2+1) = \mathfrak{R}'_8$$

which is found in §7, and set $P = x$ and we get

$$x^4(x+1)^6(x^3+x+1)(x^3+x^2+1)(x^4+x^3+x^2+x+1) = \mathfrak{R}'_{10}.$$

Now substitute $x + 1$ for x in $\mathfrak{R}'_1$, $\mathfrak{R}'_3$, $\mathfrak{R}'_5$, $\mathfrak{R}'_8$ and $\mathfrak{R}'_{10}$ and five other one-rings are found; these will be called $\mathfrak{R}'_2$, $\mathfrak{R}'_4$, $\mathfrak{R}'_6$, $\mathfrak{R}'_9$, $\mathfrak{R}'_{11}$.

We may not begin with $x^7(x+1)^7$ or any other of the trivial type and secure new one-rings as before because of Lemma 2, which states that $x^2 + x + 1$ and $x^4 + x^3 + x^2 + x + 1$ are the only complete irreducible polynomials.

Although not proved it seems likely that these eleven one-rings are the only ones of the non-trivial type which exist.

7. **One-rings** $x^h(x+1)^kP^n$. We now find all one-rings of the form $x^h(x+1)^kP^n$. This problem will be considered in four parts:

(a) $A = x^h(x+1)^kP^{2n}$,
(b) $A = x^{2h}(x+1)^{2k}P^{2n-1}$,
(c) $A = x^{2h-1}(x+1)^{2k-1}P^{2n-1}$,
(d) $A = x^{2h}(x+1)^{2k-1}P^{2n-1}$.

(a) If A is an $\mathfrak{R}'$ in this case we have

$$x^h(x+1)^kP^{2n} \Rightarrow (x^h + \cdots + 1)[(x+1)^k + \cdots + 1]\sigma P^{2n}.$$

This implies $x^h(x+1)^k = (x^h + \cdots + 1)[(x+1)^k + \cdots + 1]$ which in turn requires $P^{2n} = \sigma P^{2n}$ which is impossible.

In case (b) we have

$$x^{2h}(x+1)^{2k}P^{2n-1} \Rightarrow (x^{2h} + \cdots + 1)[(x+1)^{2k} + \cdots + 1] \\ (P+1)(P^{n-1} + \cdots + 1)^2.$$

This requires $x^{2h} + \cdots + 1 = P = (x+1)^{2k} + \cdots + 1$ and then P^{2n-1} divides the left member and P^2 the right member, an obvious contradiction.

In case (c) we have

$$x^{2h-1}(x+1)^{2k-1}P^{2n-1} \Rightarrow (x+1)(x^{h-1} + \cdots + 1)^2 \\ x[(x+1)^{k-1} + \cdots + 1]^2(P+1)(P^{n-1} + \cdots + 1)^2.$$

Since the greatest power of P dividing the right member is even, the above equation is impossible.

Finally in case (d) we have

$$\begin{aligned} x^{2h}(x+1)^{2k-1}P^{2n-1} &\Rightarrow (x^{2h} + \cdots + 1)x \\ &[(x+1)^{k-1} + \cdots + 1]^2(P+1)(P^{n-1} + \cdots + 1)^2. \end{aligned} \tag{7.1}$$

This requires $P = x^{2h} + \cdots + x + 1$ and $P + 1 = x(x + 1)(x^{h-1} + \cdots + 1)^2$. Divide (7.1) by $x^2(x + 1)P$, then extract square root of each member and we have

$$x^{h-1}(x + 1)^{k-1}P^{n-1} = [(x + 1)^{k-1} + \cdots + 1] \cdot (x^{h-1} + \cdots + 1)(P^{n-1} + \cdots + 1). \tag{7.2}$$

From (7.2) we have $P^{n-1} + \cdots + 1 = x^{\alpha}(x + 1)^{\beta}$, hence $n \neq 3, 5, 7, 9, \cdots$. We also see that P^{n-1} equals or divides $(x + 1)^{k-1} + \cdots + 1$. The only other possible factor being x we know $n \neq 4, 6, 8, 10, \cdots$. Combining results, we know n must be 1 or 2.

If $n = 2$, $P = (x + 1)^{k-1} + \cdots + 1 = x^{2h} + \cdots + 1$ and

$$x^{2h}(x + 1)^{2k-1}P^3 \Rightarrow P \cdot x \cdot P^2(P + 1)^3,$$
$$x^{2h}(x + 1)^{2k-1} = x(P + 1)^3.$$

Since $x^2 \nmid P + 1$, $2h = 4$ and $k = 5$, and we should have

$$P = x^4 + x^3 + x^2 + x + 1 = (x + 1)^4 + (x + 1)^3 + (x + 1)^2 + (x + 1) + 1. \tag{7.3}$$

However, (7.3) is not true and $n \neq 2$. If $n = 1$, (7.1) becomes

$$x^{2h}(x + 1)^{2k-1}P \Rightarrow (x^{2h} + \cdots + 1)[(x + 1)^{2k-1} + \cdots + 1]\sigma P. \tag{7.4}$$

Here

$$x^{2h} + \cdots + 1 = P, \qquad \sigma P = x(x + 1)(x^{h-1} + \cdots + 1)^2.$$

Hence

$$x^{2h-2}(x + 1)^{2k-2} = (x^{h-1} + \cdots + 1)^2[(x + 1)^{k-1} + \cdots + 1]^2.$$

Hence $h - 1 = k - 1 = 2^n - 1$, $2h = 2k = 2^n$.

By Lemma 2 the only possible values are $2h = 2k = 2$ or 4. These values give us

$$A = x^2(x + 1)(x^2 + x + 1) = \mathfrak{R}_1', \tag{7.5}$$

$$A = x^4(x + 1)^3(x^4 + x^3 + x^2 + x + 1) = \mathfrak{R}_5'. \tag{7.6}$$

Substitution of $x + 1$ for x in (7.5) and (7.6) gives

$$A = x(x + 1)^2(x^2 + x + 1) = \mathfrak{R}_2', \tag{7.7}$$

$$A = x^3(x + 1)^4(x^4 + x^3 + 1) = \mathfrak{R}_6'. \tag{7.8}$$

These are four of the eleven one-rings found in §6.

This completes the proof of the following

THEOREM 9. *The only one-rings of the form* $x^h(x + 1)^kP^n$ *are* $\mathfrak{R}_1'$, $\mathfrak{R}_5'$, $\mathfrak{R}_2'$, $\mathfrak{R}_6'$ *listed above.*

We next find all one-rings of the form $x^a(x+1)^b P^c Q^d$. The problem will be considered in five parts.

(a) $A = x^h(x+1)^k P^{2m} Q^{2n}$,
(b) $A = x^h(x+1)^k P^{2m} Q^{2n-1}$,
(c) $A = x^{2h}(x+1)^{2k} P^{2m-1} Q^{2n-1}$,
(d) $A = x^{2h}(x+1)^{2k-1} P^{2m-1} Q^{2n-1}$,
(e) $A = x^{2h-1}(x+1)^{2k-1} P^{2m-1} Q^{2n-1}$.

It is proved that no one-rings exist in cases (a) and (e). Case (b) yields

$$x^2(x+1)(x^2+x+1)^2(x^4+x+1) = \mathfrak{N}_3', \tag{7.9}$$

$$x(x+1)^2(x^2+x+1)^2(x^4+x+1) = \mathfrak{N}_4'. \tag{7.10}$$

Case (c) yields

$$x^4(x+1)^4(x^4+x^3+x^2+x+1)(x^4+x^3+1) = \mathfrak{N}_7'. \tag{7.11}$$

Case (d) yields

$$x^6(x+1)^3(x^3+x^2+1)(x^3+x+1) = \mathfrak{N}_8', \tag{7.12}$$

$$x^3(x+1)^6(x^3+x+1)(x^3+x^2+1) = \mathfrak{N}_9'. \tag{7.13}$$

The study of case (e) leads to

THEOREM 10. *There are no one-rings of the form*

$$A = x^{2h-1}(x+1)^{2k-1}\Pi P^{2m-1}.$$

We may now state the following

THEOREM 11. *The only one-rings of the form $x^a(x+1)^b P^c Q^d$ are the following:* $\mathfrak{N}_3'$, $\mathfrak{N}_4'$, $\mathfrak{N}_7'$, $\mathfrak{N}_8'$, $\mathfrak{N}_9'$.

Next we show that the only one-rings of the form $x^h(x+1)^k PQR$ are $\mathfrak{N}_{10}'$, $\mathfrak{N}_{11}'$. We begin by finding all one-rings of the form

$$x^{2h}(x+1)^{2k}PQR \Rightarrow (x^{2h}+\cdots+1)\,[(x+1)^{2k}+\cdots+1]\sigma P\sigma Q\sigma R. \tag{7.14}$$

Let $x^{2h}+\cdots+1 = P$ and $(x+1)^{2k}+\cdots+1 = Q$, then $\sigma P = x(x+1)C_1^2$ and $\sigma Q = x(x+1)C_2^2$ where C_1 and C_2 are complete polynomials. Dividing out common factors, we see that equation (7.14) becomes

$$x^{2h-2}(x+1)^{2k-2}R = C_1^2C_2^2\sigma R, \tag{7.15}$$

which is impossible since R may not divide the right unless R^2 does so.

Next let $P = (x+1)^{2k}+\cdots+1$ and $QR = x^{2h}+\cdots+1$. Then $\sigma P = (x+1)x^{2k-1}$ and $2k-1 = 2^n-1$. Hence $2k = 2$ or 4 by Lemma 2. Now

$2k = 2$ is impossible since $(x+1)^3$ divides the right member of (7.14). Set $2k = 4$, divide out the common factors in (7.14) and we get

$$x^{2h-3}(x+1)^3 = \sigma Q \sigma R. \tag{7.16}$$

Now $Q \cdot R = x^{2h} + \cdots + 1$ and by Theorem 8 we know $2h = 6$ and $Q = x^3 + x^2 + 1$, $R = x^3 + x + 1$.

Inserting these values in (7.14), we obtain the one-rings $\mathfrak{R}'_{10}$ and $\mathfrak{R}'_{11}$, previously mentioned in §6.

This proves

THEOREM 12. *The only one-rings of the form* $x^{2h}(x+1)^{2k}PQR$ *are* $\mathfrak{R}'_{10}$ *and* $\mathfrak{R}'_{11}$.

In a similar manner, we may prove the following three theorems.

THEOREM 13. *There are no one-rings* $x^{2h}(x+1)^{2k-1}PQR$.

THEOREM 14. *There are no one-rings of the form* $x^{2h}(x+1)^{2k}\Pi P^{2^n}$.

THEOREM 15. *There are no one-rings of the form* $x^{2h}(x+1)^{2k}\Pi P^{2^n-1}$ *for all* $n > 1$.

THEOREM 16. *Given* $x^h(x+1)^k \prod_{i=1}^{n} P_i = \mathfrak{R}'$, *where* $\mathfrak{R}'$ *is a one-ring having no factors of this form which are one-rings, then* $\mathfrak{R}'$ *is uniquely determined by* h *and* k.

By Theorem 10 we know at least one exponent h or k is even. To construct the one-ring, h and k being given, we first set

$$x^h(x+1)^k \rightarrow x^a(x+1)^b P_1 P_2 \cdots P_k. \tag{7.17}$$

Insert each P_i appearing on the right in (7.17), into the left member and then each σP_i on the right and again each new P_i on the right, into the left and continue this process until no new irreducible factors appear when σP_i is formed, i.e., until $\sigma P_i = x^d(x+1)^e$.

We have then placed identical irreducible polynomials on each side. On the left we have $x^h(x+1)^k$ and on the right perhaps $x^\alpha(x+1)^\beta$ but since the two members are of the same degree in x we have $h + k = \alpha + \beta$. We know, however, that $\mathfrak{R}'$ is completed and $h = \alpha$, $k = \beta$ because if $\mathfrak{R}'$ were incomplete and we inserted a missing factor P_m on the left of our ring under construction and σP_m on the right, we would have $\alpha + \beta > h + k$ which is impossible.

8. One-rings which are perfect squares.

THEOREM 17. *There are no one-rings*

$$A = P^{2m}Q^{2n}.$$

Proof. $P^{2m}Q^{2n} \rightrightarrows (P^{2m} + \cdots + 1)(Q^{2n} + \cdots + 1)$ implies $P^{2m} = Q^{2n} + \cdots + 1$ and $Q^{2n} = P^{2m} + \cdots + 1$ both of which are impossible by Lemma 5.

THEOREM 18. *There are no one-rings ΠP_i^2, where all P_i are of the same degree.*

Proof. If there were such a one-ring we would have

$$P_k^2 + P_k + 1 = P_1 P_m,$$
$$P_j^2 + P_j + 1 = P_1 P_n.$$

Adding these equations we have

$$(P_k + P_j)(P_k + P_j + 1) = P_1(P_m + P_n).$$

This equation implies $P_1 = P_k + P_j$ or $P_1 = P_k + P_j + 1$ both of which are impossible because the right members are of lower degree than P_1.

THEOREM 19. *There are no one-rings ΠP_i^4, where all P_i are of the same degree.*

Proof. By Lemmas 5 and 6, in such a one-ring an irreducible T would appear in four equations such as

$$P^4 + P^3 + P^2 + P + 1 = TA, \tag{8.1}$$
$$Q^4 + Q^3 + Q^2 + Q + 1 = TB, \tag{8.2}$$
$$R^4 + R^3 + R^2 + R + 1 = TC, \tag{8.3}$$
$$S^4 + S^3 + S^2 + S + 1 = TD. \tag{8.4}$$

From these we get

$$\begin{aligned}(P + Q + R + S + 1)(P + Q)(P + R)(P + S)(Q + R)(Q + S)(R + S) \\ = T[(A + B)(P + R)(P + S)(R + S) \\ + (A + C)(P + Q)(P + S)(Q + S) \\ + (A + D)(P + Q)(Q + R)(P + R)].\end{aligned} \tag{8.5}$$

This equation implies that T must equal one of the factors on the left, but if P, Q, R, S, T are of the same degree this is impossible.

The next two theorems are stated without proof.

THEOREM 20. *If $\Pi P^{2^n} = \mathfrak{R}'$ is a one-ring such that no factor of the form A^{2^n} is a one-ring it is uniquely determined by any irreducible factor P. Consequently no P appears with the same exponent in any two such one-rings.*

THEOREM 21. *If ΠP_i^2 $(i = 1, 2, \cdots, n)$ is a one-ring, for all i we have $P_i = A_i(x^2 + x + 1) + 1$ and P_i is of even degree.*

9. **Two-rings,** $x^\alpha(x + 1)^\beta \leftrightarrow A$. Corresponding to the infinite set of trivial one-rings is the infinite set of two-rings,

$$x^{2^n-1}(x + 1)^{2^m-1} \leftrightarrow x^{2^m-1}(x + 1)^{2^n-1} \qquad (m \neq n). \tag{9.1}$$

Preliminary to finding all $\mathfrak{R}^2$ having one member $x^\alpha(x+1)^\beta$ we consider the solutions of

$$x^\alpha(x+1)^\beta \to A \to x^\gamma(x+1)^\delta. \tag{9.2}$$

We first note some facts.

If P^m / A but $P^{m+1} \nmid A$ then $m = 2^n - 1$. By Lemma 2 and Theorem 8, if either α or β is even the only permissible values are 2, 4, and 6.

If $\alpha = 2^{\alpha_1} - 1$ then $P \nmid \sigma x^\alpha$.

If $\alpha = 2\alpha_1$ then $P / \sigma x^{2\alpha_1}$ but $P^2 \nmid \sigma x^{2\alpha_1}$.

If $\alpha = 2\alpha_1 - 1 \neq 2^{\alpha_2} - 1$ then the highest power of P dividing $\sigma x^{2\alpha_1} - 1$ is P^{2^r}.

If P / A we have $2^n - 1 = 2^r + 1$ so that $r = 1$, $n = 2$ or $r = 0$, $n = 1$, that is, only P or P^3 may divide A. Except in the trivial $\mathfrak{R}^2$, both α and β may not be odd.

The only combinations of even values of α and β which satisfy (9.2) are the following:

(a) $x^2(x+1)^4 \to (x^2+x+1)(x^4+x^3+1) \to x^4(x+1)^2$,
(b) $x^2(x+1)^6 \to (x^2+x+1)(x^3+x^2+1)(x^3+x+1) \to x^4(x+1)^4$,
(c) $x^4(x+1)^4 \to (x^4+x^3+x^2+x+1)(x^4+x^3+1) \to x^4(x+1)^4$,
(d) $x^4(x+1)^6 \to (x^4+x^3+x^2+x+1)(x^3+x^2+1)$
$(x^3+x+1) \to x^4(x+1)^6$,

and three others obtained by substituting $x+1$ for x in (a), (b), and (d).

If, of α and β, one is even and the other of the form $2^n - 1$ then only P'/A. If $\beta = 2, 4, 6$, we have the following solutions:

$$x^{2^n-1}(x+1)^2 \to (x+1)^{2^n-1}(x^2+x+1) \to x^{2^n}(x+1),$$
$$x^{2^n-1}(x+1)^4 \to (x+1)^{2^n-1}(x^4+x^3+1) \to x^{2^n+2}(x+1),$$
$$x^{2^n-1}(x+1)^6 \to (x+1)^{2^n-1}(x^3+x^2+1)(x^3+x+1) \to x^{2^n+2}(x+1)^3.$$

Other solutions are constructed by substituting $x+1$ for x in each of the three sets of solutions.

If P^3 / A, one of the exponents α and β is even (2, 4, 6) and the other is odd (5, 9, 13). The (4, 9) combination does not satisfy (9.2). The other combinations do satisfy (9.2) and give the following:

$$x^5(x+1)^2 \to (x+1)(x^2+x+1)^3 \to x^4(x+1)^3,$$
$$x^2(x+1)^5 \to x(x^2+x+1)^3 \to x^3(x+1)^4,$$
$$x^{13}(x+1)^6 \to (x+1)(x^3+x^2+1)^3(x^3+x+1)^3 \to x^{10}(x+1)^9,$$
$$x^6(x+1)^{13} \to x(x^3+x+1)^3(x^3+x^2+1)^3 \to x^9(x+1)^{10}.$$

All two-rings having one member $x^\alpha(x+1)^\beta$ may be found among the solutions of (9.2).

We may now state

Theorem 22. *The only two-rings having one member of the form $x^\alpha(x+1)^\beta$ are the following*:

$$\mathfrak{R}_1^2\colon x^4(x+1)^4 \leftrightarrow (x^4+x^3+x^2+x+1)(x^4+x^3+1),$$
$$\mathfrak{R}_2^2\colon x^4(x+1)^6 \leftrightarrow (x^4+x^3+x^2+x+1)(x^3+x^2+1)(x^3+x+1),$$
$$\mathfrak{R}_3^2\colon x^6(x+1)^4 \leftrightarrow (x^3+x^2+1)(x^3+x+1)(x^4+x^3+1).$$

This result may be extended to find all $\mathfrak{R}^{2^n}$ of the form

$$x^{\alpha_1}(x+1)^{\beta_1} \to A_1 \to x^{\alpha_2}(x+1)^{\beta_2} \to A_2 \to \cdots \to A_n \to x^{\alpha_1}(x+1)^{\beta_1}. \tag{9.3}$$

They are found among the solutions of (9.2) and we may therefore state

THEOREM 23. *The complete list of $\mathfrak{R}^{2^n}$ of the form (9.3) consists of the three $\mathfrak{R}^2$ above and the following six $\mathfrak{R}^4$:*

$$\begin{aligned}
\mathfrak{R}^4_1\colon\ & x^2(x+1)^4 \to (x^2+x+1)(x^4+x^3+1) \to x^4(x+1)^2 \\
&\to (x^4+x^3+x^2+x+1)(x^2+x+1) \to x^2(x+1)^4, \\
\mathfrak{R}^4_2\colon\ & x(x+1)^2 \to (x+1)(x^2+x+1) \to x^2(x+1) \\
&\to x(x^2+x+1) \to x(x+1)^2, \\
\mathfrak{R}^4_3\colon\ & x^4(x+1) \to x(x^4+x^3+x^2+x+1) \to x(x+1)^4 \\
&\to (x+1)(x^4+x^3+1) \to x^4(x+1), \\
\mathfrak{R}^4_4\colon\ & x^3(x+1)^4 \to (x+1)^3(x^4+x^3+1) \to x^6(x+1) \\
&\to x(x^3+x^2+1)(x^3+x+1) \to x^3(x+1)^4, \\
\mathfrak{R}^4_5\colon\ & x^4(x+1)^3 \to x^3(x^4+x^3+x^2+x+1) \to x(x+1)^6 \\
&\to (x+1)(x^3+x^2+1)(x^3+x+1) \to x^4(x+1)^3, \\
\mathfrak{R}^4_6\colon\ & x^3(x+1)^6 \to (x+1)^3(x^3+x^2+1)(x^3+x+1) \\
&\to x^6(x+1)^3 \to x^3(x^3+x^2+1)(x^3+x+1) \to x^3(x+1)^6.
\end{aligned}$$

The following theorems are proved by similar methods:

THEOREM 24. *The only two-rings of the form*

$$x^h(x+1)^kP \leftrightarrow x^a(x+1)^bA, \qquad a > 0,\ b > 0,$$

are

$$x^6(x+1)^3(x^4+x^3+x^2+x+1) \leftrightarrow x^4(x+1)^3(x^3+x^2+1)(x^3+x+1)$$

and

$$x^3(x+1)^6(x^4+x^3+1) \leftrightarrow x^3(x+1)^4(x^3+x^2+1)(x^3+x+1).$$

THEOREM 25. *There are no $\mathfrak{R}^2$ of the form*

$$x^h(x+1)^kP^{2^n} \leftrightarrow x^aA, \qquad a \neq 0, \quad (x+1)\nmid A.$$

THEOREM 26. *There are no $\mathfrak{R}^2$ of the form*

$$x^h(x+1)^kP^{2m-1} \leftrightarrow x^a(x+1)^bA,$$

where neither a nor b is zero.

THEOREM 27. *There are no $\mathfrak{R}^2$ of the form*

$$x^h(x+1)^kP^{2m-1} \leftrightarrow x^a(x+1)^bA \qquad (m > 1).$$

THEOREM 28. *There are no $\mathfrak{R}^2$ of the form*

$$x^h(x+1)^kP\cdot Q \leftrightarrow x^a(x+1)^bR\cdot S,$$

for $h > 0$, $k > 0$, $a > 0$, $b > 0$.

MEREDITH COLLEGE.

THE CANONICAL LINES

By Chuan-Chih Hsiung

1. Introduction. Associated with a point P of a surface S several covariant lines have been defined by various authors in independent investigations of the projective differential geometry of surfaces. Among them perhaps the most important are the directrices of Wilczynski,[1] the axes of Čech,[2] the edges of Green,[3] and the projective normal of Green and Fubini.[4] In view of the fact that all of these lines just mentioned that pass through the point P are characterized by apparently unrelated properties, it has been considered remarkable that they all should lie in a plane. This plane is called the *canonical plane.* Any line passing through the point P and lying in the canonical plane is spoken of as a *canonical line of the first kind.* The reciprocal polar lines of the canonical lines of the first kind with respect to the quadric of Lie or any quadric of the pencil of Darboux at the point P of the surface S, dually, lie in the tangent plane of S at P and pass through a common point, which is called the *canonical point.* Any of these lines is spoken of as a *canonical line of the second kind.*

The purpose of this note is to present a new geometric characterization for a general canonical line of each kind, and especially for the first axis of Čech.

2. Analytic basis. In ordinary space in which a point has projective homogeneous coördinates $x^{(1)}, \cdots, x^{(4)}$, the parametric vector equation of an analytic non-ruled surface S is

$$x = x(u, v), \tag{1}$$

the parameters being u, v. If the asymptotic curves on the surface S are the parametric curves, the coördinates x satisfy a system of two partial differential equations which can be reduced to *Fubini's canonical form*

$$\begin{cases} x_{uu} = px + \theta_u x_u + \beta x_v, \\ x_{vv} = qx + \gamma x_u + \theta_v x_v \end{cases} \qquad (\theta = \log \beta\gamma), \tag{2}$$

subscripts indicating partial differentiation and the coefficients being functions of u, v which satisfy certain integrability conditions. Then the coördinates of

Received March 4, 1941.

[1] E. J. Wilczynski, *Projective differential geometry of curved surfaces* (Memoirs 2-3), Transactions of the American Mathematical Society, vol. 9(1908), pp. 79–120; 293–315.

[2] E. Čech, *L'intorno di un punto d'una superficie considerato dal punto di vista proiettivo,* Annali di Matematica Pura ed Applicata, (3), vol. 31(1922), pp. 191–206.

[3] G. M. Green, *Memoir on the general theory of surfaces and rectilinear congruences,* Transactions of the American Mathematical Society, vol. 20(1919), pp. 79–153.

[4] G. Fubini, *Fondamenti della geometria proiettivo-differenziale di una superficie,* Reale Accademia delle Scienze, Torino, Atti, vol. 53(1918), pp. 1032–1043. See also G. M. Green, Bulletin of the American Mathematical Society, vol. 23(1916), pp. 73–74, Abstract.

any point in the space can be represented by the coördinates of the form $x_1 x + x_2 x_u + x_3 x_v + x_4 x_{uv}$.

At a point P_x of the surface S the curves of Darboux and the curves of Segre are respectively represented by the differential equations

$$\frac{dv}{du} = -\epsilon^i \lambda, \tag{3}$$

$$\frac{dv}{du} = \epsilon^i \lambda, \tag{4}$$

where $i = 1, 2, 3$; $\epsilon^3 = 1$, $\lambda = (\beta/\gamma)^{\frac{1}{3}}$.

The canonical line $l_1(k)$ of the first kind is the intersection of the two planes

$$x_2 - k\psi x_4 = 0, \qquad x_3 - k\phi x_4 = 0, \tag{5}$$

and the canonical line $l_2(k)$ of the second kind crosses the asymptotic tangents at the points

$$(k\phi, 1, 0, 0), \qquad (k\psi, 0, 1, 0), \tag{6}$$

where k is a constant and ϕ, ψ are defined by

$$\phi = (\log \beta\gamma^2)_u, \qquad \psi = (\log \beta^2\gamma)_v. \tag{7}$$

The osculating plane π_i of a general curve of Darboux defined by equation (3) at a point P_x of the surface S is determined by x, x', x'', where

$$\begin{cases} x' = x_u - \epsilon^i \lambda x_v, \\ x'' = x_{uu} - 2\epsilon^i \lambda x_{uv} + \epsilon^{2i}\lambda^2 x_{vv} + \epsilon^i(\epsilon^i \lambda\lambda_v - \lambda_u)x_v. \end{cases} \tag{8}$$

The local equation of this plane π_i is, by a simple calculation, found to be

$$\epsilon^i \lambda x_2 + x_3 + \left(\epsilon^{2i}\frac{\beta}{\lambda} + \epsilon^i \frac{\lambda}{3}\psi + \tfrac{1}{3}\phi\right)x_4 = 0. \tag{9}$$

Similarly, the local equations of the osculating planes of the curves of Segre defined by equations (4) at P_x of S are

$$\epsilon^i \lambda x_2 - x_3 + \tfrac{1}{3}(\epsilon^i \lambda \psi + \phi)x_4 = 0 \qquad (i = 1, 2, 3). \tag{10}$$

3. The canonical lines of the second kind. For the purpose of finding a geometric characterization of a general canonical line of the second kind, we consider a one-parameter family of polarities C_h, which was constructed by Professor Su,[5] with respect to the pencil Q_h of quadrics

$$(1 - 2h)x_2 x_3 = \{2(1 - h)x_1 + \tfrac{1}{4}\phi x_2 + \tfrac{1}{4}\psi x_3 + h_4 x_4\}x_4, \tag{11}$$

where h_4 denotes another parameter.

[5] B. Su, *The canonical edges of Green*, Tôhoku Mathematical Journal, vol. 39(1934), pp. 269–278.

From (9) and (11) it is easily seen that the poles of the osculating planes π_i of the curves of Darboux with respect to the polarity C_h are

$$(12)\quad \begin{cases} x_1 = -\dfrac{1}{24(1-h)}\left[12(1-2h)\epsilon^{2i}\dfrac{\beta}{\lambda} + (7-8h)\epsilon^i\lambda\psi + (7-8h)\phi\right], \\ x_2 = 1, \quad x_3 = \epsilon^i\lambda, \quad x_4 = 0 \qquad (i = 1, 2, 3). \end{cases}$$

The harmonic polar of the point P_x with respect to the triangle (12) is the line of the equations

$$(13)\qquad x_4 = 0, \qquad x_1 + \frac{7-8h}{24(1-h)}\phi x_2 + \frac{7-8h}{24(1-h)}\psi x_3 = 0,$$

h being a constant. These are, however, the equations of a general canonical line of the second kind. We may, therefore, state the

THEOREM. *Let V_i $(i = 1, 2, 3)$ be the poles of the osculating planes π_i of the curves of Darboux with respect to the polarity C_h at P_x of S, then the harmonic polar of P_x with respect to the triangle $V_1V_2V_3$ is a canonical line of the second kind which depends upon the polarity C_h. By a proper selection of the constant h this line may be made to become any desired canonical line of the second kind.*

In particular, when $h = \infty$, C_h is the polarity of Lie, and the line (13) is the second axis of Čech. When $h = 0$, C_h is the correspondence of Palozzi,[6] and the line (13) is the canonical line[7] $l_2(-\frac{7}{24})$. When $h = \frac{1}{4}$, C_h is the polarity with respect to any principal quadric of Lane,[8] and the line (13) is a new canonical line $l_2(-\frac{5}{18})$. When $h = \frac{5}{8}$, C_h is the correspondence which the author[9] has independently defined, the line (13) in question is also a new canonical line $l_2(-\frac{2}{9})$. When $h = \frac{7}{8}$, we have the second projective normal of Green and Fubini; etc.

4. **The canonical lines of the first kind.** Let us now consider the conic sections K_i^h of any quadric of the pencil Q_h made by the osculating planes π_i of the curves of Darboux. The equations of K_i^h are readily found to be (9) and

$$(14)\quad \begin{aligned} &(1-2h)\epsilon^i\lambda x_2^2 + 2(1-h)x_1x_4 + \left[(1-2h)\epsilon^{2i}\frac{\beta}{\lambda} + \tfrac{1}{3}(\tfrac{1}{4}-2h)\epsilon^i\lambda\psi \right. \\ &\qquad \left. + \tfrac{1}{3}(\tfrac{7}{4}-2h)\phi\right]x_2x_4 + x_4^2(\cdots) = 0 \qquad (i = 1, 2, 3). \end{aligned}$$

[6] G. Palozzi, *Alcuni risultati di geometria proiettivo-differenziale*, Rendiconti dei Lincei, (6), vol. 15(1932), pp. 543–548.

[7] For another definition of this canonical line see G. Sannia, *Nuove definizioni del fascio canonico*, Rendiconti dei Lincei, (6), vol. 8(1928), pp. 373–375.

[8] E. P. Lane, *Invariants of intersection of two curves on a surface*, American Journal of Mathematics, vol. 54(1932), pp. 699–706.

[9] C. C. Hsiung, *Note on the intersection of two space curves*, to be published in the Tôhoku Mathematical Journal.

On the other hand, the plane determined by an arbitrary canonical line $l_1(k)$ and the u-tangent intersects the planes π_i in the lines μ_i of the equations (9) and

$$(15) \qquad \epsilon^i\lambda x_2 + \left[\epsilon^{2i}\frac{\beta}{\lambda} + \epsilon^i\frac{\lambda}{3}\psi + (k + \tfrac{1}{3})\phi\right]x_4 = 0 \qquad (i = 1, 2, 3).$$

For a fixed value of h, we have three poles of μ_i with respect to the conics K_i^h on the tangent plane at P_x of S. The harmonic polar of P_x with respect to these three poles is given by the equations

$$(16) \quad x_4 = 0, \qquad x_1 - \frac{1}{24(1-h)}[24k(1-2h) + 1 - 8h]\phi x_2 + \frac{7-8h}{24(1-h)}\psi x_3 = 0.$$

For the sake of convenience, we shall call this line (16) the *associated harmonic polar* p_1 of μ_i with respect to the planes π_i and the pencil Q_h.

Interchanging ϕ, ψ and x_2, x_3 we obtain the equations of the associated harmonic polar p_2 of ν_i, which are the intersections of the planes π_i and the plane determined by $l_1(k)$ and the v-tangent:

$$(17) \quad x_4 = 0, \qquad x_1 + \frac{7-8h}{24(1-h)}\phi x_2 - \frac{1}{24(1-h)}[24k(1-2h) + 1 - 8h]\psi x_3 = 0.$$

The join of the intersection of p_1 with the v-tangent and that of p_2 with the u-tangent is the general canonical line of the second kind (13). On the contrary, the join of the intersection of p_1 with the u-tangent and that of p_2 with the v-tangent is the reciprocal $l_2(k)$ of $l_1(k)$ with respect to the surface S when, and only when,

$$(18) \qquad k = \frac{1-8h}{24h}.$$

If the osculating planes (10) of the curves (4) of Segre are used in place of the planes π_i, we may obtain the same results. Thus we arrive at the

THEOREM. *Let π_i $(i = 1, 2, 3)$ be the osculating planes of the curves of Darboux (or Segre) at P_x of S, and K_i^h the conic sections of π_i with any quadric of the pencil Q_h. Let p_1 be the associated harmonic polar of the intersections of π_i with the plane determined by an arbitrary canonical line l_1 of the first kind and an asymptotic tangent a_1 of S at P_x with respect to the planes π_i and the pencil Q_h. Similarly, let p_2 be the associated harmonic polar of the intersections of π_i with the plane determined by l_1 and the other asymptotic tangent a_2 with respect to π_i and Q_h. Then the join of the intersections of p_1 with a_2 and of p_2 with a_1 is a general canonical line $l_2((8h-7)/[24(1-h)])$, which is independent of the choice of the canonical line l_1. Moreover, the join of the intersections of p_1 with a_1 and of p_2 with a_2 is the reciprocal l_2 of l_1 with respect to S when and only when l_1 is the canonical line $l_1((1-8h)/24h)$, which may be made to become any desired canonical line of the first kind by a proper selection of the constant h.*

In particular, when $h = \infty, 0, \frac{1}{4}$ and $\frac{1}{8}$, the line $l_1((1 - 8h)/24h)$ becomes the first axis of Čech, the first canonical tangent, the principal line $l_1(-\frac{1}{6})$ of Fubini and the projective normal of Green and Fubini respectively. When $h = \frac{5}{8}$, we have a new canonical line $l_1(-\frac{4}{15})$.

It is worth noticing that *the polar line of the tangent plane at P_x of S with respect to the trihedron formed by the three osculating planes of the curves of Darboux is the first axis of Čech.*

University of Chekiang,
Kweichow, China.

A PRINCIPLE OF JESSEN AND GENERAL FUBINI THEOREMS

By Nelson Dunford and J. D. Tamarkin

Much of Jessen's fundamental work on direct product spaces, in particular the existence of measure, has in recent years been given a fairly general form. However, there is one important principle discovered by Jessen that has not been announced in the abstract form it deserves, and we propose in this note to take a step in that direction. The principle referred to is embodied in the theorem asserting that a real measurable function $f(x)$ on a direct product space, which is independent of any change of a finite number of the coördinates of x, is necessarily almost everywhere a constant. This theorem has many applications, but unfortunately the Euclidean methods of proof used by Jessen cannot be extended very far. We shall show here that the same theorem is valid even if the measure has its values in a Hausdorff ring with only two idempotents and the function f has its values in a separable metric space. We also develop the Fubini-Jessen theorems for vector-valued functions.

Notation and terminology. The letter A will be used for a class of elements α. B will always mean a non-void subset of A and σ a non-void finite subset of A. The symbols $\tilde{B}$, $\tilde{\sigma}$ will be used for the complements of B, σ in A. The symbol x/B will stand for a function defined on B. For each $\alpha \in A$ there is a space S_α and a Borel field $\mathfrak{F}_\alpha$ of sets in S_α and $S_\alpha \in \mathfrak{F}_\alpha$. We shall use the logical notation of Kuratowski, *Topologie* I, and define the direct product of sets $E_\alpha \in S_\alpha$ by

$$\mathop{P}_{\alpha \in B} E_\alpha \equiv \mathop{\&}_{x/B} \prod_{\alpha \in B} [x(\alpha) \in E_\alpha].$$

We shall sometimes write S^B in place of $\mathop{P}\limits_{\alpha \in B} S_\alpha$, and here and elsewhere the superscript B will be omitted in case $B = A$, so that whenever B and $\tilde{B}$ appear as superscripts in an equation we know that $\tilde{B} \neq 0$. Since x/B and $y/\tilde{B}$ determine uniquely a function $z = (x, y)$ on A by the convention $z(\alpha) = x(\alpha)$ or $y(\alpha)$ according as $\alpha \in B$ or $\alpha \in \tilde{B}$ and conversely z/A determines uniquely x/B and $y/\tilde{B}$, we may and shall sometimes write $S = S^B \times S^{\tilde{B}}$. An *elementary set in* S^B is by definition one of the form

$$E = \mathop{P}_{\alpha \in \sigma} E_\alpha \times S^{\tilde{\sigma} B},$$

where $E_\alpha \in \mathfrak{F}_\alpha$ and $\sigma \in B$. The Borel field in S^B determined by the elementary sets in S^B will be denoted by $\mathfrak{F}^B$ and we write $\mathfrak{F}$ in place of $\mathfrak{F}^A$. For a set $M \subseteq S$ the projection of M on S^B is defined as

$$\mathrm{Proj}_B\, M \equiv \mathop{\&}_{x/B} \sum_{y/\tilde{B}} [(x, y) \in M].$$

Received May 8, 1941.

A set $N \subseteq S$ is said to be a cylinder set over S^B in case

$$N = \mathrm{Proj}_B\, N \times S^{\bar B}. \tag{1}$$

It is clear that an elementary set in S is cylinder over some S^σ, and thus the Borel field in S determined by those sets in $\mathfrak{F}$ which are cylinder over some S^σ coincides with $\mathfrak{F}$. A set $M \subseteq S$ is said to have the property J (the Jessen property) in case it is cylinder over $S^{\bar\sigma}$ for every $\sigma \subseteq A$. In symbols, M has the property J if and only if

$$M = \mathrm{Proj}_{\bar\sigma}\, M \times S^{\sigma}, \qquad \sigma \subseteq A. \tag{2}$$

Another way of stating this is to say that M has the property J if and only if every pair $x, y \in S$ with $x(\alpha) = y(\alpha)$ except for α in some finite set σ must both belong to M or to $S - M$. A function f/S with values in any set is said to have the property J (the Jessen property) in case for every pair $x, y \in S$ with $x(\alpha) = y(\alpha)$ except for α in some finite set σ we have $f(x) = f(y)$. Thus a set $M \subseteq S$ has the property J if and only if its characteristic function does.

Basic assumptions. We shall be concerned with a Hausdorff ring R with a unit 1. About R we assume that the ring operations $a + b$, $a \cdot b$, $a - b$ are continuous functions of the two variables. For each $\alpha \in A$ let Φ_α be a completely additive set function on $\mathfrak{F}_\alpha$ to R with $\Phi_\alpha(S_\alpha) = 1$. The property of complete additivity is

$$\Phi_\alpha\Big(\sum_{n=1}^{\infty} E_n\Big) = \sum_{n=1}^{\infty} \Phi_\alpha(E_n), \qquad E_n \in F_\alpha,\ E_n E_m = 0,\ n \neq m.$$

Since the left side of this equation is independent of the order of arrangement of the terms E_n, so is the right side. We are now in a position to state the basic concept underlying the Jessen theory. For each $B \subseteq A$ we assume the existence of a completely additive set function Φ^B on $\mathfrak{F}^B$ to R satisfying the three conditions

$$\Phi^B = \Phi_\alpha \qquad \text{if } B = \alpha, \tag{3}$$

$$\Phi(E^B \times E^{\bar B}) = \Phi^B(E^B) \cdot \Phi^{\bar B}(E^{\bar B});\ E^B \in \mathfrak{F}^B,\ E^{\bar B} \in \mathfrak{F}^{\bar B}, \tag{4}$$

$$\Phi^B(S^B) = 1. \tag{5}$$

Here we have written Φ for Φ^A. It follows from these conditions that

$$\Phi^B\Big[\mathop{P}_{\alpha \in \sigma} E_\alpha \times S^{\bar\sigma B}\Big] = \prod_{\alpha \in \sigma} \Phi_\alpha(E_\alpha); \qquad \sigma \in B,\ E_\alpha \in \mathfrak{F}_\alpha. \tag{6}$$

Although the ring R is not assumed to be commutative, the product on the right side of the above equation is commutative since the left side is independent of order.

In the case where R is the real number system the existence of Φ^B has been

established under restrictions on A, S_α, and Φ_α by Jessen [2][1] and in complete generality by Kolmogorov [4], von Neumann [5] and Doob [1]. The results of these mathematicians may be used to establish readily its existence for certain other rings R, e.g., the ring, with the weak topology of Tychonoff, of real functions over an abstract set. We shall not discuss here the question of existence.

Some theorems of Jessen.

THEOREM 1. *If M, $N \in \mathfrak{F}$ and M has the property* J, *then*

$$\Phi(MN) = \Phi(M)\cdot\Phi(N). \tag{7}$$

It is clear that the family of all N which satisfy (7) is a Borel field containing S. Thus, in view of the remark following (1), it suffices to prove (7) in case N is cylinder over some S^σ. Note first that by (1), (4) and (5) we have for any cylinder set N

$$\Phi(N) = \Phi^B(\mathrm{Proj}_B\, N), \qquad N \text{ cylinder over } S^B. \tag{8}$$

If M is the set (2) and N the set (1) with $B = \sigma$, we have

$$MN = \mathrm{Proj}_{\bar\sigma}\, M \times \mathrm{Proj}_\sigma\, N,$$

and so by (4) and (8)

$$\Phi(MN) = \Phi^{\bar\sigma}(\mathrm{Proj}_{\bar\sigma}\, M)\cdot\Phi^\sigma(\mathrm{Proj}_\sigma\, N) = \Phi(M)\cdot\Phi(N).$$

COROLLARY 1. *If each one of a sequence M_n of sets in $\mathfrak{F}$ has the property* J, *then*

$$\Phi\left(\prod_{n=1}^{\infty} M_n\right) = \lim_m \prod_{n=1}^{m} \Phi(M_n). \tag{9}$$

COROLLARY 2. *If $M \in \mathfrak{F}$ has the property* J, *then $\Phi(M)$ is idempotent.*

A function f on S to a topological space E is said to be measurable in case $f^{-1}(D) \in \mathfrak{F}$ for every Borel set $D \subseteq E$.

THEOREM 2. *Let f be a measurable function on S to a separable metric space E. If f has the property* J *and R has no idempotents other than* 0 *and* 1, *then $f(x)$ is a constant except on a set where $\Phi = 0$.*

For every Borel set $D \subseteq E$, the set $f^{-1}(D)$ has the property J; so by Corollary 2, $\Phi[f^{-1}(D)]$ is 0 or 1. Since E is separable and $\Phi[f^{-1}(E)] = 1$, there is for every $n = 1, 2, \cdots$ a Borel set D_n of diameter less than $1/n$ with $\Phi[f^{-1}(D_n)] = 1$. Thus by Corollary 1,

$$\Phi\left[f^{-1}\left(\prod_{n=1}^{\infty} D_n\right)\right] = \Phi\left[\prod_{n=1}^{\infty} f^{-1}(D_n)\right] = \lim_m \prod_{n=1}^{m} \Phi[f^{-1}(D_n)] = 1.$$

Since diam $\prod_{n=1}^{\infty} D_n = 0$, the theorem is proved.

[1] Numbers in brackets refer to the references at the end of the paper.

The ring of all continuous real functions on a connected space is an example of a ring with 0, 1 as its only idempotents. In the theorem, if it is assumed only that 1 is isolated among the idempotents, we have been able to prove nothing except that f has almost all its values in a totally bounded analytic set.

The Fubini-Jessen theorems. Here we shall be concerned with functions on S^B to a Banach space E. Such a function is said to be elementary in case it is constant over each of a finite number of elementary sets in S^B and zero elsewhere. The function is said to be strongly measurable in case it is the limit almost everywhere of a sequence of elementary functions. Such a function has almost all its values in a separable subset of E, and hence Theorem 2 is applicable regardless of the character of E. We take R to be the real number system and for simplicity assume that Φ^B is non-negative. Thus by (4) the measures Φ^B may be interpreted as a set of independent probability distributions. The space $L_p^B(E)$, $1 \leqq p < \infty$, by definition consists of those strongly measurable functions f on S^B to E for which

$$\|f\|_p^B \equiv \left(\int_{S^B} \|f(x)\|^p d\Phi^B\right)^{1/p} < \infty.$$

In the above notation we omit the B in case $B = A$, and we omit E in case it is the real number system. The finite sets $\sigma \subseteq A$ form a directed set and the limits $\lim_\sigma F_\sigma$ that appear in what follows are always taken in the sense of E. H. Moore. That is, $\lim_\sigma F_\sigma = F$ if and only if

$$\prod_{\epsilon>0} \sum_{\sigma} \prod_{\sigma' \supset \sigma} [\|F_{\sigma'} - F\| < \epsilon].$$

LEMMA 1. *The map $f^B = T^B f$ defined by*[2]

$$f^B(x) = \int_{S^{\bar{B}}} f(x)\, d\Phi^{\bar{B}}$$

is bounded and linear with norm 1 when regarded as either mapping $L_p(E)$ onto itself or onto $L_p^B(E)$.

It is obviously linear. It is bounded, for

$$\begin{aligned}
\int_S \|f^B(x)\|^p d\Phi &= \int_{S^{\bar{B}}} d\Phi^{\bar{B}} \int_{S^B} \|f^B(x)\|^p d\Phi^B \\
&= \int_{S^B} \|f^B(x)\|^p d\Phi^B \leqq \int_{S^B} d\Phi^B \int_{S^{\bar{B}}} \|f(x)\|^p d\Phi^{\bar{B}} \qquad (10)\\
&= \int_S \|f(x)\|^p d\Phi.
\end{aligned}$$

[2] The integral may be taken as that defined by S. Bochner, N. Dunford, G. Birkhoff, or B. J. Pettis as all of these integrals coincide over the class of functions we are considering.

Here we have used the Fubini theorem over the product $S^{\beta} \times S^{\bar\beta}$. The inequality follows from the fact that the L_1 norm of a real function is no greater than its L_p norm.

THEOREM 3. *For $f \in L_p(E)$ we have*

$$\lim_{\sigma} f^{\sigma} = f \qquad \textit{in } L_p(E).$$

For a function which is constant on a given elementary set in S the theorem is obvious. Since this type of function forms a fundamental set in $L_p(E)$, the theorem is an immediate consequence of the lemma.

THEOREM 4. *For $f \in L_p(E)$ we have*

$$\lim_{\sigma} f^{\bar\sigma} = \int_S f(x)\, d\Phi \qquad \textit{in } L_p(E).$$

Let $y = \int_S f(x)\, d\Phi$, then

$$\begin{aligned} f^{\bar\sigma}(x) - y &= \int_{S^{\sigma}} f(x)\, d\Phi^{\sigma} - \int_{S^{\sigma}} d\Phi^{\sigma} \int_{S^{\bar\sigma}} f(x)\, d\Phi^{\bar\sigma} \\ &= \int_{S^{\sigma}} [f(x) - f^{\sigma}(x)]\, d\Phi^{\sigma}. \end{aligned}$$

If in (10) we replace $f^{\beta}(x)$ by $f^{\bar\sigma}(x) - y = [f(x) - y]^{\bar\sigma}$, we get

$$\begin{aligned} \int_S \| f^{\bar\sigma}(x) - y \|^p\, d\Phi &= \int_{S^{\bar\sigma}} \| f^{\bar\sigma}(x) - y \|^p\, d\Phi^{\bar\sigma} \\ &\leqq \int_{S^{\bar\sigma}} d\Phi^{\bar\sigma} \int_{S^{\sigma}} \| f(x) - f^{\sigma}(x) \|^p\, d\Phi^{\sigma} \\ &= \int_S \| f(x) - f^{\sigma}(x) \|^p\, d\Phi, \end{aligned}$$

and so the theorem follows from Theorem 3.

LEMMA 2.[3] *Let f be strongly measurable on S to E and $\sigma_1 \subset \sigma_2 \subset \cdots$. Let $g_n(x)$ be either $\| f(x) \|^{\sigma_n}$ or $\| f(x) \|^{\bar\sigma_n}$. Let $g(x) = \underset{1 \leqq n < \infty}{\text{l.u.b.}}\, g_n(x)$, $\delta > 0$, $S_\delta = \underset{x}{\mathfrak{S}}[g(x) > \delta]$. Then*

$$\delta \Phi(S_\delta) \leqq \int_{S_\delta} \| f(x) \|\, d\Phi.$$

Suppose $g_n(x) = \| f(x) \|^{\bar\sigma_n}$. Then the set

$$S_k \equiv \underset{x}{\mathfrak{S}}\, [\underset{1 \leqq n \leqq k}{\text{l.u.b.}}\, g_n(x) > \delta]$$

[3] This lemma is due to Jessen [2]. The proof given here is due essentially to I. Halperin.

is cylinder over S^{σ_k} (it does not depend on the coördinates $x(\alpha)$ with $\alpha \in \tilde{\sigma}_k$) and $S_k \uparrow S_\delta$. Let

$$T_1 = S_1, \qquad T_k = S_k - S_{k-1}, \qquad\qquad k > 1,$$

so that $S_\delta = \sum_{k=1}^{\infty} T_k$ is a disjoint sum and T_k is cylinder over S^{σ_k}. Thus by (1) and (8)

$$T_k = \mathrm{Proj}_{\sigma_k} T_k \times S^{\tilde{\sigma}_k}, \qquad \Phi^{\sigma_k}(\mathrm{Proj}_{\sigma_k} T_k) = \Phi(T_k),$$

and hence the Fubini theorem gives

$$\int_{T_k} \| f(x) \| \, d\Phi = \int_{\mathrm{Proj}_{\sigma_k} T_k} d\Phi^{\sigma_k} \int_{S^{\tilde{\sigma}_k}} \| f(x) \| \, d\Phi^{\tilde{\sigma}_k}$$
$$= \int_{\mathrm{Proj}_{\sigma_k} T_k} g_k(x) \, d\Phi^{\sigma_k} \geqq \delta\Phi(T_k).$$

Summing on k gives the desired result. A similar method applies when $g_n(x) = \| f(x) \|^{\sigma_n}$.

The lemma enables us to replace the convergence asserted in Theorems 3 and 4 by pointwise convergence in the case where A is denumerable.[4]

THEOREM 5. *Suppose A is the denumerable set $\{\alpha_n\}$ and $\sigma_n = (\alpha_1, \cdots, \alpha_n)$. Then for $f \in L(E)$ we have*

$$\lim_n f^{\tilde{\sigma}_n}(x) = \int_S f(x) \, d\Phi \qquad \textit{almost everywhere on } S.$$

Let $y = \int_S f(x) \, d\Phi$ and $\delta(x) = \limsup_n \| f^{\tilde{\sigma}_n}(x) - y \|$. The function $\delta(x)$ is a constant δ almost everywhere since it is measurable and has the property J (Theorem 3). We shall make an indirect proof by assuming $\delta > 0$. Let $h(x) = \| f(x) - g(x) \|$, where g is an elementary function and $\int_S h(x) \, d\Phi < \frac{1}{4}\delta$. For large n we have $g^{\tilde{\sigma}_n}(x) = \int_S g(x) \, d\Phi$ and hence for large n

$$\| f^{\tilde{\sigma}_n}(x) - y \| - \tfrac{1}{4}\delta \leqq \| f^{\tilde{\sigma}_n}(x) - \int_S g(x) \, d\Phi \|$$
$$= \left\| \int_{S^{\sigma_n}} (f(x) - g(x)) \, d\Phi^{\sigma_n} \right\| \leqq h^{\tilde{\sigma}_n}(x).$$

Hence for almost all $x \in S$, $\mathrm{l.u.b.}_n \, h^{\tilde{\sigma}_n} \geqq \frac{3}{4}\delta$. An application of Lemma 2 gives

$$\tfrac{3}{4}\delta \leqq \int_S h(x) \, d\Phi \leqq \tfrac{1}{4}\delta,$$

which is a contradiction to the assumption that $\delta > 0$.

[4] This was done for real functions with restricted S_α, Φ_α by B. Jessen [2] and was stated, without proof, for real functions and unrestricted S_α, Φ_α by B. Jessen and A. Wintner [5].

THEOREM 6. *Suppose A is the denumerable set $\{\alpha_n\}$ and $\sigma_n = (\alpha_1, \cdots, \alpha_n)$. Then for $f \in L(E)$ we have*

$$\lim_n f^{\sigma_n}(x) = f(x) \qquad \textit{almost everywhere on } S.$$

In view of Theorem 3 it is sufficient to show that the $\lim_n f^{\sigma_n}(x)$ exists almost everywhere. Let

$$\psi(x) \equiv \lim_{q\to\infty} \operatorname*{l.u.b.}_{m,n>q} \| f^{\sigma_n}(x) - f^{\sigma_m}(x) \|.$$

Let $h(x) = f(x) - g(x)$, where g is elementary and $\int_S \| h(x) \| \, d\Phi < \epsilon$. For large n, $g^{\sigma_n}(x) = g(x)$, and thus for large n, m

$$\begin{aligned} \| f^{\sigma_n}(x) - f^{\sigma_m}(x) \| &\leqq \| f^{\sigma_n}(x) - g(x) \| + \| f^{\sigma_m}(x) - g(x) \| \\ &= \| h^{\sigma_n}(x) \| + \| h^{\sigma_m}(x) \|. \end{aligned}$$

Hence

$$\psi(x) \leqq 2 \operatorname*{l.u.b.}_n \| h^{\sigma_n}(x) \| \leqq 2 \operatorname*{l.u.b.}_n \| h(x) \|^{\sigma_n}.$$

Thus

$$\mathop{\mathfrak{E}}_x [\psi(x) > 2\delta] \subseteq \mathop{\mathfrak{E}}_x [\operatorname*{l.u.b.}_n \| h(x) \|^{\sigma_n} > \delta],$$

and Lemma 2 gives

$$\delta\Phi\{\mathop{\mathfrak{E}}_x [\psi(x) > 2\delta]\} \leqq \int_S \| h(x) \| \, d\Phi < \epsilon.$$

Since δ, ϵ are both arbitrary positive numbers $\Phi\{\mathop{\mathfrak{E}}_x [\psi(x) > 0]\} = 0$ and the proof is complete.

REFERENCES

1. J. L. DOOB, *Stochastic processes with an integral-valued parameter*, Transactions of the American Mathematical Society, vol. 44(1938), pp. 87–150.
2. B. JESSEN, *The theory of integration in a space of an infinite number of dimensions*, Acta Mathematica, vol. 63(1934), pp. 249–323.
3. B. JESSEN AND AUREL WINTNER, *Distribution functions and the Riemann zeta function*, Transactions of the American Mathematical Society, vol. 38(1935), pp. 48–88.
4. A. KOLMOGOROV, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Mathematik, vol. 2(1933), no. 3.
5. J. VON NEUMANN, *Functional operators*, Lectures delivered at the Institute for Advanced Study, 1933–1934.

YALE UNIVERSITY AND BROWN UNIVERSITY.

DIVISION ALGEBRAS OVER A FUNCTION FIELD

By A. A. Albert

1. **Introduction.** Let F be the centrum of a simple algebra A and L be a subfield of F. Then we shall call a field W over L a *splitting field over*[1] F *of* A if a composite[2] of W and F, regarded as a scalar extension of F, splits A. If x is an indeterminate over L and F is algebraic of finite degree over the rational function field $L(x)$, we shall call such a field W a *constant splitting field over* F of D. It is known[3] that the algebraic closure of L is such a splitting field, and it follows[4] that there exists a constant splitting field over F of finite degree over L for any A.

The most important splitting fields K of a division algebra D of degree m over its centrum F are those fields of least possible degree over F. This degree is m and it is known[5] also that all such fields are equivalent to subfields of D of maximal degree over F. It is natural then to propose the question as to the universal existence of a constant splitting field over F of degree m over L for any D. In particular a cyclic field of this kind would provide an extremely simple generation of D.

We shall answer the question just proposed in the negative and shall prove indeed that the least degree n of a constant splitting field may be arbitrarily large. We shall prove in fact the

Theorem 1. *Let x be an indeterminate over a finite field L, $F = L(x)$, m and n be positive integers. Then there exist division algebras D of degree m over F as centrum and with n as the least degree over L of a constant splitting field over F of D if and only if m divides n and every prime factor of n divides m.*

Our proof of this theorem depends fundamentally upon the hypothesis that L is finite, and that we should make such a restriction is quite natural in a first study of the problem we have proposed. However, the conclusions of our theorem are probably valid under very different and varied hypotheses on L, and even when L is a non-modular field. As an indication in this direction we shall prove

Received June 27, 1941; presented to the American Mathematical Society Sept. 2, 1941.

[1] If the degree of F over L is finite the algebra D is a simple (division) algebra over L and a field W over L is generally called a splitting field of D if the scalar extension D_W is a *direct sum* of total matric algebras, the number of them being the degree of F over L. This is *not* the concept in which we are interested and so we insert the phrase *over F*.

[2] We shall be interested usually in the case where W is normal of finite degree over L, and in this case all composites over F of W and F are equivalent. Then we may and shall speak of *the* composite of W and F. See the author's *Modern Higher Algebra*, Sec. 8.4.

[3] C. C. Tsen, *Divisions algebren über Funktionenkörpern*, Gött. Nach., 1933, pp. 335–9.

[4] This follows from the author's *Structure of Algebras*, Sec. 11.2.

[5] Ibid., Theorem 4.27.

THEOREM 2. *Let L be a field such that there exist at least two inequivalent quadratic fields over L. Then there is a division algebra D of degree two over its centrum $L(x)$ such that $W(x)$ does not split D for any quadratic field W of degree two over L.*

The case in this theorem of algebras of degree two over a field of characteristic two leads one to a consideration of cyclic p-algebras, that is, cyclic algebras of degree p^e over a centrum F of characteristic p. We shall study particularly the case of algebras of degree p. Each such algebra is uniquely defined by two quantities c and g in F, and we shall write[6] $A = (c, g]$. If $F = L(x)$, and c may be taken to be in L, then A has a splitting field $W(x)$, where W is cyclic of degree p over L. It is known[7] that if L is perfect we may always take $g = x$, $A = (c, x]$, and it is natural then to make the error of assuming that it is possible to choose c to be a polynomial of $L[x]$. However, we shall show that if $A = (c, x]$ and c is in $L[x]$ we may take $A = (d, x]$ for d in L. Then, if in addition L is finite, all such algebras may be seen to be equivalent. But there exist algebras $(d, x - a]$, for a in L, which are not equivalent to $(d, x]$.

The importance of the study of cyclic p-algebras of degree p for the problem of constant splitting fields of p-algebras of degree p^e is enhanced by a device which we shall give reducing the proofs of existence theorems for the latter algebras to those for the former. We shall use a modification of this device with Theorem 2 and shall derive

THEOREM 3. *Let there exist at least two inequivalent separable quadratic fields over a field L of characteristic two. Then for any integer $e \geqq 1$ there exists a cyclic division algebra D of degree 2^e over its centrum $L(x)$ with no splitting field $W(x)$, for W of degree 2^e over L.*

It is reasonable that our existence theorems should all employ cyclic algebras, and this is particularly true in the case of p-algebras. For all such algebras are cyclically representable, that is, similar[8] to cyclic algebras. It was thought likely,[9] at one time, that all cyclically representable algebras are cyclic, and the author subsequently[10] set up the mechanism necessary to determine whether or not all 2-algebras of degree 4 are cyclic. We may now announce the unpublished result of a master's dissertation of June 1940 by Louis Gordon at the University of Chicago, in which this mechanism was used to show that there do

[6] This notation is used by E. Witt in his *Zyklische Körper und Algebren der Charakteristik p vom Grad p^n*, Journal für Mathematik, vol. 176(1937), pp. 126–140. Some of the auxiliary lemmas we shall obtain are included in the theorems of Witt, but it would require more discussion to obtain our precise results as instances of those of Witt than to derive them. Incidentally, the results we shall use all have their beginnings in the author's original investigations on p-algebras as given in *Structure of Algebras*, Chapter VII.

[7] Cf. the author's *p-algebras over a field generated by one indeterminate*, Bulletin of the American Mathematical Society, vol. 43(1937), pp. 733–6.

[8] *Structure of Algebras*, Theorem 7.31.

[9] This conjecture was made by H. Hasse in a letter to the author written in 1931.

[10] See the author's *Normal division algebras of degree four over F of characteristic two*, American Journal of Mathematics, vol. 56(1934), pp. 75–86.

exist non-cyclic algebras of degree and exponent four[11] over $L(x, y)$ where L is any infinite field of characteristic two, and x and y are independent indeterminates over L.

Gordon's result implies that for $n > 2$ if L is finite of characteristic two and $F = L(x_1, \cdots, x_r)$ for independent indeterminates $x_1, \cdots, x_r$ over L, then there exist non-cyclic algebras of degree four over F. Let then D be a division algebra of degree four over its centrum K, which is algebraic of finite degree over F, and suppose $r \leqq 2$. If $r = 1$, it is well known that D is a cyclic algebra. We shall conclude our study with a brief proof that also if $r = 2$ all division algebras D of degree four over K are cyclic.

2. Rational function fields over a finite field. Our principal existence theorem is a consequence of the valuation theory of any field $F = L(x)$ of rational functions $a = a(x)$ in an indeterminate x with coefficients in a finite field L. We shall give a résumé of that theory here. The field F is the quotient field of the integral domain $J = L[x]$ of all polynomials $a(x)$, and all the ideals of J are principal ideals. All algebraic extensions W of finite degree over L are cyclic fields[12] over L and thus are the (necessarily normal) stem fields of equations $g(x) = 0$, for irreducible polynomials $g = g(x)$ of J. But the prime ideals of J are the principal ideals (g), for g irreducible, and W may be constructed as the difference ring $J - (g)$. Hence W is uniquely determined by its degree, and this is the degree of $g(x)$.

If $a \neq 0$ is in F we may write $a = g^e bc^{-1}$, where b and c are in J and are not divisible by g, and

$$e = e_g(a)$$

is an integer. Write $e_g(0) = \infty$ and thus define what is generally called a *value group* for F and the given g. Then there is a corresponding *valuation* $\| a \| = 2^{-e}$, and we may construct[13] the completion of F with respect to this valuation. This is a field F_g with the same value group as F and, in our present case, may be chosen so that each quantity of F has the same value $e(a)$ as it had originally in F. We have thus extended the function $e(a)$ to be a function on F_g to the value group of F.

The quantities a of F_g for which $e(a) \geqq 0$ form an integral domain J_g whose quotient field is F_g. The ideals of J_g are all principal and are indeed powers of its only prime ideal, the principal ideal (g) of all quantities a of J_g for which $e(a) > 0$. Thus every a of J_g has the form

$$a = g^{e(a)}b,$$

where b is a unit of J_g. Indeed the units of J_g consist of all its quantities b for which $e(b) = 0$. The difference ring $J_g - (g)$ is equivalent to $J - (g)$ and thus is a finite field whose degree r over L is the degree of the polynomial g.

[11] Algebras of degree four and exponent two are necessarily cyclic algebras.
[12] Cf. the author's *Modern Higher Algebra*, sec. 7.10.
[13] Ibid., sec. 11.11.

The finiteness of $J_g - (g)$ may be seen to imply the validity of the Hensel Lemma[14] on p-adic fields and consequently[15] that F_g contains a primitive $(q^r - 1)$-th root of unity z. Then $L(z)$ has degree r over L and is equivalent to $J_g - (g)$. $L(z)$ is the root field of $g(x) = 0$. It is the maximal finite subfield of F_g in the strong sense that it contains all finite subfields of F_g. Assume henceforth that q is the number of elements in L.

Let us suppose at this point that W is any field of degree n over L. Then W is a cyclic field over L, $W = L(z_n)$ for a primitive $(q^n - 1)$-th root of unity z_n. The correspondence $z_n \to z_n^S = z_n^q$ then induces a generating automorphism of the Galois group of W over L as well as a generating automorphism of the cyclic field $Z = W(x) = F(z_n) = L(x, z_n)$ over F. Now Z is normal over F and hence has a single composite (in the sense of equivalence over F_0) with any extension field F_0 over F. Thus we form the composite Z_g of Z and F_g. We now prove

LEMMA 1. *Let d be the greatest common divisor of $n = bd$ and the degree r of an irreducible $g = g(x)$ in $L(x)$, Z_g and F_g be defined as above. Then $Z_g = F_g(z_n)$ is cyclic of degree b over F_g.*

For the composite of Z and F_g is the composite of W and F_g and hence is $F_g(z_n)$. The intersection W_0 of W and F_g is a finite field, and hence is the intersection[16] of W and the maximal finite $L(z)$ of F_g. Since there is one and only one finite field of a given degree over L, the degree of W_0 is the largest common divisor of n and r. That Z_g is cyclic of degree b over F_g now follows from the general theory of composites.

The field Z_g is the completion[17] of Z with respect to valuations of Z preserving our given valuation of F, and its maximal finite field is $L(z_n)$. Thus Z_g is what is called an *unramified* field over F_g and a quantity a of F_g is a norm

$$a = N_{Z_g/F_g}(b) \qquad (b \text{ in } Z_g),$$

if and only if $e(a)$ is divisible by the degree of Z_g over F_g. In particular all units of J_g are such norms, the cyclic algebras (Z_g, S, u) are total matric algebras for all units u of J_g, and every (Z_g, S, a) is equivalent[18] to $(Z_g, S, g^{e(a)})$.

We now continue our study of the valuations of $F = L(x)$. Our field $F = L(y)$, for $y = x^{-1}$, and F is the quotient field of $\mathfrak{J} = L[y]$ as well as of $L[x]$. Then every irreducible polynomial of $L[y]$ determines a valuation of F. However, if

$$a = a(x) = a_0 + a_1 x + \cdots + a_t x^t$$

for a_i in L and $a_0 \neq 0$, $a_t \neq 0$, we may write

$$a = (a_0 y^t + \cdots + a_t) y^{-t} = \bar{a} y^{-t},$$

[14] The proof in *Modern Higher Algebra*, sec. 12.10, is valid.

[15] Cf. M. Deuring, *Algebren*, p. 137.

[16] *Modern Higher Algebra*, sec. 8.4.

[17] Ibid., sec. 11.12.

[18] For these properties of cyclic algebras see *Structure of Algebras*, sec. 5.9.

where $\bar{a}$ is in $\bar{J}$ and is irreducible if and only if a is irreducible. If g is an irreducible polynomial of J, and a in J has the form $a = g^e b$ for b not divisible by g, then $\bar{a}y^{-t} = \bar{g}^e \bar{b} y^{-s}$. Hence $\bar{g}$ determines the same valuation as does g unless $\bar{g} = y$. However, it is clear that $e_y(x) = -1$ and that $e_g(x) \geqq 0$ for every irreducible g of J. Thus the only valuation of F obtained from the consideration of $L[y]$ and not already obtained is that determined by $g(y) = y$. The degree of this polynomial in y is unity and since the valuation theory is as before we may interpret Lemma 1 to give

LEMMA 2. *Let F_y be the completion of $F = L(x)$ with respect to the valuation defined by $g(y) = y$ in $L[y]$, $y = x^{-1}$. Then the composite Z_y of F_y and a field W of degree n over L is cyclic of degree n over F_y.*

The only valuations of $L(x)$ are those described above and it is also known[19] that the index m of a simple algebra A over its centrum F is the least common multiple of the indices of the scalar extensions A_{F_y}, A_{F_g} for all the irreducible polynomials g of $L[x]$. We shall now proceed to apply this theory to construct an algebra A of degree n and index m and thus a proof of the result we desire.

3. **The principal existence theorem.** If $W = L(z_n)$ is a field of degree n over L, the field $Z = W(x) = F(z_n) = L(x, z_n)$ is cyclic of degree n over $F = L(x)$ with generating automorphism S over F induced by $z_n^S = z_n^q$. Thus we may define a cyclic algebra[20]

$$A = (Z, S, h)$$

for h in F. If m is the index of A it divides the degree n of A and $n = mr$ for an integer r. We now select h to be an irreducible polynomial of degree r in $L[x]$.

To compute the index of A we see first that if g is an irreducible polynomial of $L[x]$ then

$$A_{F_g} \sim (Z_g, S, h).$$

If $g \neq h$ the quantity h is a unit of J_g, h is a norm in Z_g over F_g, A_{F_g} has index unity. If $h = g$, Lemma 1 implies that m is the degree of Z_h over F_h. The least power of h which is a norm in Z_h over F_h is then m and (Z_g, S, h) has exponent m and must be a division algebra. A_{F_h} has index m. Now $h = h(x) = h(y)y^{-r}$ so that

$$A_{F_y} = (Z_y, S, h) = (Z_y, S, y^{-r}).$$

But then A_{F_y} is similar to $(Z_y, S, y^{-1})^r$. The algebra (Z_y, S, y) has degree and index n and is reciprocal to (Z_y, S, y^{-1}) whose r-th direct power then has index m.

[19] These results are given in E. Witt, *Riemann-Rochscher Satz und ζ-Funktion im Hyperkomplexen*, Math. Ann., vol. 110(1934), pp. 12–28.

[20] The results on algebras used in the remainder of this section are to be found in the author's *Structure of Algebras*.

We have thus shown that A has index m, $A = M \times D$ where M is a total matric algebra and D is a division algebra of degree m over its centrum F. Clearly W is a constant splitting field over F of A and hence of D. The least degree j of such a field is at most n.

Let us now proceed to a proof of our main theorem. We assume that $n = mr$, where every prime factor of r divides m, and construct D as above. We now let N be a field of the least degree j over L such that $N(x)$ splits $D, j \leqq n$. The degree of $N(x)$ over F is j and hence m divides $j = sm$, $s \leqq r$. If $s < r$ there is a prime factor p of r such that

$$m = p^e m_0, \qquad r = p^f r_0, \qquad s = p^t s_0,$$

for integers m_0, r_0, s_0 prime to p and

$$e > 0, \qquad f > t \geqq 0.$$

Then the greatest common divisor d of r and $j = p^{f+t} m_0 s_0$ is divisible by p^{t+1} so that $j = bd$ where b is not divisible by p^f. Hence b is not divisible by m. However, if $K = N(x)$ splits D, the field K_h must split A_{F_h} of index m over F_h, and m must divide the degree of K_h over F_h. This degree is b by Lemma 1, a contradiction. Hence $j = n$.

Conversely, if D is a division algebra of degree m over $L(x)$ and N has degree $n = n_1 n_2$ over L for n_2 prime to n_1, then $N = N_1 \times N_2$ where N_i has degree n_i over L. If n_2 is prime to M then $N(x)$ splits D if and only if $N_1(x)$ splits D. Hence, N cannot be minimal unless every prime factor of its degree divides m. This completes our proof.

It would be desirable to extend our existence theorem to the case where the centrum F of D is an algebraic extension of finite degree over $L(x)$. The valuation theory permits a rather complete analysis of this case but requires a more comprehensive study of the existence of algebraic extensions of L whose composite with F has prescribed ramification order and residue class degree than has yet been made. We therefore leave this question as an unsolved problem in our theory.

4. **Algebras of degree two.** We know now that a division algebra of degree m over its centrum $L(x)$ need not have any splitting field $W(x)$, for W an algebraic extension of degree m over L, at least in the case where L is a finite field. As we stated in our introduction it is desirable to see that this property is not a function of the finiteness of L and that it may hold, in particular, when L is non-modular. We shall thus give an indication in this direction by proving Theorem 2.

Let us assume first that the characteristic of L is not two so that there exist quantities a and b in L such that $a \neq k^2$, $b \neq k^2$. $b \neq k^2 a$ for any k of L. For every $h = h(x)$ of $L[x]$ we write $h(k)$ for the quantity of L obtained by replacing x by k in L and prove

LEMMA 3. *Let $c = a - x$, $g = x(x + b - a)$, d be any quantity of L. Then*

$$f = cu^2 - dv^2 + (w^2 - ct^2)g = 0$$

for u, v, w, t in $L(x)$ only if $u = v = w = t = 0$.

For if $f = 0$ for u, v, w, t not all zero we may assume that these quantities are all in $L[x]$ and that their g.c.d. is unity. If u and v are divisible by x, then x^2 divides $(w^2 - ct^2)g$ and x divides $w^2 - ct^2$. If x divides t, then x divides w, contrary to hypothesis. Thus $[w(0)]^2 - a[t(0)]^2 = 0$ for $t(0) \neq 0$, and a is the square of a quantity of L, contrary to hypothesis.

We see now that x divides $cu^2 - dv^2$ but does not divide v. It follows, as in the proof above, that $d = k^2a$ for $k \neq 0$ in L. Now $x + b - a$ divides $cu^2 - dv^2$, and thus $b[u(a - b)]^2 - k^2a[v(a - b)]^2 = 0$. This is not possible, since b is not the product of a by the square of a quantity of L, unless $v(a - b) = 0$ and $x + b - a$ divides both u and v. Then $x + b - a$ divides $w^2 - ct^2$, $[w(a - b)]^2 - b[t(a - b)]^2 = 0$, and $x + b - a$ divides w and t. This is a contradiction.

We use the quantities c and g given in Lemma 3 to define an algebra $D = (1, i, j, ij)$ over $F = L(x)$ such that $i^2 = c$, $j^2 = g$, $ji = -ij$. Then D is a simple algebra with F as centrum. A quantity z in D is not in F yet is such that $z^2 = d$ in F if and only if $z = ui + wj + tij$ for u, w, and t not all zero and in F, $d = cu^2 + gw^2 - cgt^2$. Then D is a total matric algebra if and only if some such z has the property $z^2 = d = 0$. This is impossible by Lemma 3. Thus D is a division algebra. If $W(x)$ splits D for W of degree two over L, we have $W = L(y)$, $W(x) = F(y)$, $y^2 = d$ in L. Then there must be a quantity z in D and not in F such that $z^2 = d$. This is also impossible, and we have completed our proof of Theorem 2 in the case where the characteristic of L is not two.

Suppose now that L has characteristic two. The hypothesis of Theorem 2 then implies the existence of quantities a and b in L such that $a \neq k^2 - k$, $b \neq k^2 - k$, $a + b \neq k^2 - k$ for any k of L. We define

$$c = x + a, \qquad g = x(x + a - b)$$

as before and note our use of $-1 = 1$. Define an algebra $D = (1, i, j, ij)$ over $F = L(x)$ such that

$$i^2 = i + c, \qquad j^2 = g, \qquad ji = (i + 1)j.$$

If k is a scalar such that $k^2 = k + c$ we may represent D as the set of all matrices

$$z = \begin{pmatrix} z_1 + z_2k & z_3g + z_4kg \\ z_3 + z_4k + z_4 & z_1 + z_2k + z_2 \end{pmatrix},$$

and see that

$$z^2 + z_2z + N(z) = 0,$$

where

$$N(z) = z_1^2 + z_1z_2 + z_2^2c + (z_3^2 + z_3z_4 + z_4^2c)g.$$

The separable subfields $F(z)$ of D are all obtained by taking $z_2 = 1$ and are equivalent to fields defined by roots of $y^2 = y + d$, where $d = N(z)$. Making the result homogeneous in the variables z_1, z_3, z_4 we are lead to consider the quadratic form

$$f = u^2 + uv + v^2(c - d) + (w^2 + wt + t^2c)g.$$

Then we will have proved that there exists no separable field W of degree two over L such that $W(x)$ splits D if we can prove that the hypothesis that d is in L implies that $f \neq 0$ for any relatively prime u, v, w, t in $L[x]$. Let us then proceed to this proof using first the fact that x is a factor of g.

If x divides both u and v, then x divides $w^2 + wt + t^2c$, so that $[w(0)]^2 + [w(0)t(0)] + a[t(0)]^2 = 0$. If $t(0) \neq 0$, then $a = k^2 - k$ where $k = w(0)[t(0)]^{-1}$, a contradiction. Hence $t(0) = 0$, $w(0) = 0$, x divides u, v, w, t. Thus x cannot divide both u and v and must not divide v. It follows that $[u(0)]^2 + u(0)v(0) + [v(0)]^2(a - d) = 0$, $d = a + k^2 - k$ for k in L. Then d defines the same quadratic extension of L as does a and we may assume, without loss of generality, that $d = a$. Note that our proof already implies that $d \neq 0$ and our algebra is not equivalent to the total matric algebra $A = (1, i, j, ij)$ with $i^2 = i$, $ji = (i + 1)j$, $j^2 = g$. Thus D is a division algebra and we are taking $c - d = x$.

We now use the fact that $x + a - b$ divides g and hence also $u^2 + uv + v^2x$. If $x + a - b$ divides u and v, then $x + a - b$ divides $w^2 + wt + t^2c$, so that $[w(b - a)]^2 + [w(b - a)][t(b - a)] + [t(b - a)]^2b = 0$ whence $t(b - a) \neq 0$ would imply that $b = y^2 - y$ for y in L contrary to hypothesis. But $t(b - a) = 0$ implies that $w(b - a) = 0$, $x + a - b$ divides u, v, w, t. Thus $x + a - b$ cannot divide both u and v and must not divide v, $v(b - a) \neq 0$. But $[u(b - a)]^2 + [u(b - a)][v(b - a)] + [v(b - a)]^2(b - a) = 0$ and thus $b - a = b + a = y^2 - y$ for $y = u(b - a)[v(b - a)]^{-1}$ in L. This final contradiction completes our proof.

It remains to show that there exist no inseparable fields W of degree two over L such that $W(x)$ splits D. This is true if no quantity z of D not in F has the property $z^2 = d$ in L. But then otherwise $z_2 = 0$, and we need to show that the equation

$$f = u^2 + dv^2 + (w^2 + wt + t^2c)g = 0,$$

for u, v, w, t in $L[x]$, is impossible if the g.c.d. of u, v, w, t is unity and d is in L, $d \neq k^2$ for k in L. By the proof above x cannot divide both u and v. But then $v(0) \neq 0$, $[u(0)]^2 + d[v(0)]^2 = 0$, $d = k^2$ for $k = u(0)[v(0)]^{-1}$ in L, a contradiction. We have completed the proof of Theorem 2.

Note that in our construction of D we have used two quantities of L such that the root field Q of the product of the corresponding quadratic equations has degree four over L. If then K is a field of degree prime to two over L, the corresponding root field over K still has degree four. But this implies that the scalar extension $D_{K(x)}$ has the property of our theorem.

5. **Some properties of p-algebras.** A generalized cyclic algebra A of degree n over F is a supplementary sum of linear subspaces[21] which may be written as

$$A = (Z, S, g) = Z + Zj + \cdots + Zj^{n-1}.$$

Here Z is a commutative separable algebra of order n over F, S is an automorphism over F of order n of Z, and

$$j^r z = z^{S^r} j^r, \qquad j^n = g \qquad (r = 0, 1, \cdots, n - 1),$$

for $g \neq 0$ in F. If Z is a field over F, the algebra A is an ordinary cyclic algebra. Otherwise Z is the direct product of a cyclic field and a diagonal algebra, and A is a total matric algebra if Z is a diagonal algebra.

Let $n = p$ be a prime, so that Z is either a field or a diagonal algebra. Suppose also that p is the characteristic of F. Then we may always assume that $Z = F[u]$ is the set of all polynomials in u and in the unity quantity of A, and that

$$u^S = u + 1, \qquad u^p = u + c,$$

for c in F. Then Z is a cyclic field or a diagonal algebra according as there does not or does exist a quantity h in F such that $c = h^p - h$. In particular, Z is a diagonal algebra if $c = 0$.

Observe now that our algebras A of degree p are uniquely determined by two quantities c and g in F where $g \neq 0$. Thus we introduce the notation

$$A = (c, g]$$

for such algebras. The usual condition that a cyclic algebra be a total matric algebra is that g be the norm of a quantity of Z. However, the norm form is a very complicated function of our parameter c, and we shall obtain a much simpler condition.

Observe first that if $g = h^p$ for h in F, the algebra A is always a total matric algebra. $(c, g] = (d, g]$ for any c and d. Hence let us assume that g is not the p-th power of any quantity of F. Then the subalgebra $F(j)$ of A defined for $j^p = g$ is a maximal subfield of A, and if y in A has the property that $y^p = g$ there is an automorphism[22] S of A such that $j = y^S$.

Let T be an equivalence over F

$$b \leftrightarrow b^T \qquad (b \text{ in } B, \quad b^T \text{ in } A)$$

of $A = (c, g]$ and $B = (d, g]$. Then B contains quantities v and y such that $v^p = v + d$, $yv = (v + 1)y$, $y^p = g$. It follows that y^T in A has the property $(y)^p = g$ and that there exists an automorphism S of A such that $y^{TS} = j$. Let $w = v^{TS}$ so that w is in A,

$$w^p = w + d, \qquad jw = (w + 1)j.$$

[21] For a consideration of such algebras see *Structure of Algebras*, Chapters VI and VII.

[22] Ibid., Theorem 4.14. The condition $g \neq h^p$ simplifies our reference but is really not essential.

But then $j(w - u) = [(w + 1) - (u + 1)]j = (w - u)j$,

$$w = u + f,$$

where $fj = jf$, and f is necessarily in $F(j) = F[j]$.

Conversely, if f is any polynomial in j with coefficients in F and $w = u + f$, we have $jf = fj$, $ju = (u + 1)j$ so that $jw = (w + 1)j$. The inner automorphism

$$S: \qquad a \leftrightarrow jaj^{-1} = a^S$$

of A has order p and induces a corresponding automorphism $w \leftrightarrow w + 1$ in the algebra V of all polynomials in 1 and w with coefficients in F. Each such polynomial $a(w) = a_0 + a_1 w + \cdots + a_t w^t$, for a_i uniquely determined in F, $a_t \neq 0$, $0 \leqq t < p$. The coefficient of w^{t-1} in a^S is $a_{t-1} + a_t t \neq a_{t-1}$ unless $t = 0$, so that $a^S \neq a$ unless a is in F. But then $(w^p - w)^S = (w + 1)^p - (w + 1) = w^p - w$ is in F.

We have now shown that

$$(1) \qquad w = u + f(j), \qquad w^p = w + c + d(f),$$

where $d(f)$ is in F. To compute $d(f)$ we observe that if y is any quantity of the field $F(j)$ which is not in F, then y^p is in F. Suppose also that A contains a quantity v for which $yv = (v + 1)y$, and that $V = F[v]$. Then

$$A = V + Vy + \cdots + Vy^{p-1}, \qquad A = (b, h],$$

where

$$v^p = v + b, \qquad y^p = h.$$

We take $w = v + y$ and our proof shows that

$$w^p = w + b + d$$

for d in F. But $w^p = (v + y)^p = v^p + y^p + a$, where a is a sum of a finite number of expressions each a product of s factors y and $p - s$ factors v for $0 < s < p$. Hence a is in $Vy + \cdots + Vy^{p-1}$. Also $w^p = v + y + b + d = v + b + h + a$, so that since $v + b + d$ and $v + b + h$ are in V we have $a = y$, and

$$(2) \qquad d = h = y^p.$$

We shall use the result just obtained to prove that if in (1) we write

$$(3) \qquad f = f(j) = f_0 + f_1 j + \cdots + f_{p-1} j^{p-1},$$

then we have

$$(4) \qquad d(f) = f_0^p - f_0 + f_1^p \cdot g + \cdots + f_{p-1}^p \cdot g^{p-1}.$$

Formula (4) holds for $f = f_0$ in F since $(u + f_0)^p = u^p + f_0^p = u + c + f_0^p = (u + f_0) + c + f_0^p - f_0$. Assume (4) for polynomials of degree at most $t - 1$

in j, so that if f has degree t, we may write $f - f_t j^t = z$, where z has degree at most $t - 1$ in j. Then the hypothesis of our induction implies that $(u + z)^p = u + z + tb$, where we are assuming that $0 < t < p$, and we have

$$tb = c + f_0^p - f_0 + f_1^p g + \cdots + f_{t-1}^p g^{t-1}.$$

Now $j(u + z) = (u + z + 1)j$, $j^t(u + z) = (u + z + t)j^t$, and if we write $v = t^{-1}(u + z)$, then $j^t v = (v + 1)j^t$. Suppose $y = t^{-1}f_t j^t$ so that

$$v^p = v + b, \qquad y^p = h = t^{-1}f_t^p g^t.$$

By (2) we have $(v + y)^p = v + y + b + h$ and $[t(v + y)]^p = t(v + y)^p = t(v + y) + tb + th$. Clearly $t(v + y) = u + f$, $th = f_t^p g^t$ and we have (4) for f of degree t. This completes our proof of (1), (3), (4).

Our result may now be stated[23] as

LEMMA 4. *Let $A = (c, g]$, $B = (d, g]$ over F. Then A and B are equivalent if and only if*

$$d = c + f_0^p - f_0 + f_1^p g + \cdots + f_{p-1}^p g^{p-1} \tag{5}$$

for f_i in F.

As an immediate corollary we have

LEMMA 5. *An algebra $(c, g]$ over F is a total matric algebra if and only if $c = f_0^p - f_0 + f_1^p g + \cdots + f_{p-1}^p g^{p-1}$ for f_i in F.*

We may use (5) to obtain a property of p-algebras connected with our study of constant splitting fields. The author has shown[24] that if L is a perfect field of characteristic p and A is a simple algebra of degree p over its centrum $F = L(x)$ then

$$A = (c, x]. \tag{6}$$

Here c is, in general, not a polynomial in x, but merely a rational function. However, as we stated in our introduction, we may prove that if $A = (c, x]$, where c is in $L[x]$, then $A = (d, x]$ for d in L.

To prove this result we observe that we are assuming that A has at least one generation (6) with c in $L[x]$. Hence we may write $A = (d, x]$ for d in $L[x]$ and having the least degree of all polynomials c occurring in the generations (6) of A for c in $L[x]$. The field L is perfect, so that the term of highest degree in x of d has the form $r^p x^t$ for $0 \leqq t < p$ and r in $L[x]$. If $t > 0$ we use (5) with $x = g$ to replace d by $d - r^p x^t$ of lower degree than d, while if $t = 0$ then we use (5) to replace d by $d - r^p + r$ which has lower degree in x than d unless r is in L. Hence, d must be in L.

We have now seen that if $A = (c, x]$ over $L(x)$, with c in $L[x]$, then A has a

[23] A result of this type was also given by N. Jacobson, *p-algebras of exponent p*, Bulletin of the American Mathematical Society, vol. 43(1937), pp. 667–70. See also O. Teichmüller, *p-Algebren*, Deutsche Mathematik, vol. 1(1936), pp. 362–388.

[24] See footnote 8.

splitting field $W(x)$ such that W is cyclic of degree p over L. However, $A = (c, x]$ may have $W(x)$ as splitting field even if c is not in $L(x)$. For example, if d is in L then $(d, x] = (c, x]$ for $c = d + f_{p-1}^p x^{p-1} = d + x^{-1}$ when we take $f_{p-1} = x^{-1}$.

One might try to show, as a converse to Theorem 3, that if A is simple of degree p over its centrum $L(x)$ and has a splitting field $W(x)$, W cyclic of degree p over L, then $A = (c, x]$ for c in $L[x]$. *This is false.* For let us take L finite so that if u and v define cyclic fields of degree p over L these fields are equivalent, $u^p = u + d$, $v^p = v + b$, necessarily $b = f_0^p - f_0 + d$ for f_0 in L. Then all algebras $(d, x]$ over $L(x)$, for d in L are cyclic algebras (Z, S, x) for the same Z and are equivalent. However, if $B = (d, g]$, the results of Section 2 imply that B_{F_h} has index unity for h irreducible in $L[x]$, B_{F_g} has index p for g irreducible in $L[x]$. In particular, the algebras $(d, x - g_0]$ defined for distinct values g_0 in L are not equivalent and no $(d, x - g_0]$ defined for $g_0 \neq 0$ has the form $(c, x]$ for c in $L[x]$.

Let us now describe a device which may be used in our study of constant splitting fields. We propose to construct division algebras D of degree p^e over a centrum $F = L(x)$, of characteristic p, and we wish to prescribe an integer n such that for every W of degree not greater than n the field $W(x)$ does not split D. Construct first a cyclic division algebra B of degree p over $L(x)$ with the required property for the given n. Then there always exists[25] a cyclic algebra D of degree p^e such that $D^t \sim B$ for $t = p^{e-1}$. But then D has exponent p^e and is a division algebra, $[D_{W(x)}]^t \sim B_{W(x)}$ is a division algebra, $D_{W(x)}$ is a division algebra and this is actually a stronger result than we require.

We shall use a modification of the device just described in our proof of Theorem 3. We construct an algebra B of degree two over $L(x)$, where L has characteristic two and B is the division algebra constructed in the proof of Theorem 2. Then it is known that there exists a cyclic algebra D of degree 2^e over $L(x)$ such that $D^m \sim B$, $m = 2^{e-1}$. Let W of degree 2^e over L have the property that $W(x)$ splits D. The field W is contained in a normal field N of degree $2^f r$ over L, for r odd, and N contains a Sylow subfield K of degree r over L such that N is metacyclic over K. The composite W_K of W and K is their direct product. It has degree 2^e over K and is metacyclic. Also $W_K(x)$ splits D and W_K contains a quadratic field S over K. By the remark at the end of §4 the algebra $B_{K(x)}$ is not split by $R(x)$ for any quadratic field R over K. Thus $B_{S(x)}$ is a division algebra similar to $[D_{S(x)}]^m$, $D_{S(x)}$ of degree 2^e has exponent 2^e and is a division algebra. This is impossible since $W_K(x)$ has degree m over $S(x)$ and splits $D_{S(x)}$.

6. **The structure of 2-algebras of degree 4.** We close our study with a remark on division algebras D of degree and exponent four over a centrum K which is algebraic of finite degree over $F = L(x, y)$. Assume that L is a finite

[25] *Structure of Algebras*, Lemma 7.11.

field and that x and y are independent indeterminates over L. We consider a quadratic form f in m variables. Then it is known[26] that f is a null form if $m > 2^{r+1}$, where $F = L(x_1, \cdots, x_r)$ for independent indeterminates $x_1, \cdots, x_r$ over L, and thus $r = 2$ in our case. Hence f is a null form precisely if $m \geqq 9$. But the condition that D be non-cyclic is exactly[27] that a certain quadratic form in nine variables be not a null form. Thus all division algebras of degree four over K of finite degree over $L(x, y)$ are cyclic algebras.

It seems reasonable that the method above may have some further application to the study of p-algebras of degrees p and p^2, but at present the author can see no obvious process of application.

THE UNIVERSITY OF CHICAGO.

[26] See footnote 5 of the author's *Quadratic null forms over a function field*, Annals of Mathematics, vol. 39(193), pp. 494–505 for this result of Chevalley, Warning, and Artin.

[27] See the reference in footnote 10.

A PROPERTY OF BANACH SPACES

By Mahlon M. Day

By a method involving a general form of integration, due to Hildebrandt, Goldstine [3] proved that a certain kind of weak completeness is necessary and sufficient for reflexivity of a Banach space B.[1] The present note gives for this theorem a new proof, based on a simple geometrical property (Lemma 1) of every Banach space. The nature of the proof suggests a new criterion (Theorem 2) for reflexivity of a Banach space; Theorem 3 collects a number of criteria for reflexivity of every λ-separable subspace of a Banach space.

If E is a subset of B, let $r(E) = \inf_{b \in E} \| b \|$ with the usual convention that if E is empty, $r(E) = +\infty$.

Lemma 1. *If $\beta_1, \cdots, \beta_k \in B^*$, if $E = \{b \mid \beta_i(b) = c_i$ for $i = 1, \cdots, k\}$ and if $M = \sup | \Sigma_{i \leq k} t_i c_i |/\| \Sigma_{i<k} t_i \beta_i \|$ where the supremum is taken[2] over all choices of the real numbers $t_1, \cdots, t_k$, then $M = r(E)$.*

This is essentially Helly's theorem; a short proof due to Mimura is quoted in Kakutani [4].

A set X of elements x is called *directed* by a relation $>$ if (1) $x_1 > x_2$ and $x_2 > x_3$ imply that $x_1 > x_3$, and (2) if x_1 and x_2 are in X there is an x_3 in X such that $x_3 > x_1$ and $x_3 > x_2$.[3] If Y is any topological space and f any function on X to Y, f converges to y or $y = \lim_x f(x)$ if and only if for each neighborhood N of y there is an x_N in X such that $f(x) \in N$ if $x > x_N$. Following Goldstine, say that B is *weakly complete relative to* X if the conditions (1) $\| b_x \| \leqq K$ for every X, and (2) $\lim_x \beta(b_x)$ exists for every β in B^*, together imply that a b in B exists for which $\lim_x \beta(b_x) = \beta(b)$ for every β in B^*. B is *weakly complete* if it is weakly complete relative to every directed set X.

The *weak neighborhoods* of a point b_0 in B are the sets

$$N = N(b_0; \beta_1, \cdots, \beta_k; \epsilon) = \{b \mid | \beta_i(b) - \beta_i(b_0) | < \epsilon \text{ for } i = 1, \cdots, k\}$$

for every choice of $\epsilon > 0$, the integer k, and the points $\beta_1, \cdots, \beta_k$ in B^*. The weak topology can be defined in the same way in a conjugate space B^* but another topology is often more useful. The *weak* neighborhoods* of β_0 in B^* are

Received February 3, 1941.

[1] It is assumed that the reader is familiar with the definition of a Banach space B, and its conjugate space B^*; see for example Banach [1]. B is reflexive if for each $\mathfrak{b}$ in B^{**}, the second conjugate space of B, there is a b in B such that $\mathfrak{b}(\beta) = \beta(b)$ for all β in B^*.

[2] $\{x \mid \cdots\}$ means the set of all x satisfying the conditions following the vertical bar. For $a_1, a_2 \geqq 0$ we shall make the convention that no matter what a_2 is, $a_1/a_2 = 0$ if $a_1 = 0$; also $a_1/a_2 = +\infty$ if $a_2 = 0$ and $a_1 > 0$.

[3] Directed sets were first studied by Moore and Smith [5]; G. Birkhoff [2] adapted this notion of convergence to topological uses.

the sets $N^* = N^*(\beta_0 ; b_1, \cdots, b_k ; \epsilon) = \{\beta \mid |\beta(b_i) - \beta_0(b_i)| < \epsilon$ for $i = 1, \cdots, k\}$ for all choices of $\epsilon > 0$, k a positive integer, and $b_1, \cdots, b_k$ in B. Let $\mathfrak{N}$ be the set of weak* neighborhoods of[4] θ in B^{**}; then $\mathfrak{N}$ is a directed set if $N > N'$ means $N \subset N'$. If $\mathfrak{N}_\mathfrak{b}$ is the set of weak* neighborhoods of $\mathfrak{b}$ in B^{**}, $\mathfrak{N}_\mathfrak{b}$ is also directed and the transformation associating $N(\theta; \beta_1, \cdots, \beta_k ; \epsilon)$ in $\mathfrak{N}$ with $N(\mathfrak{b}; \beta_1, \cdots, \beta_k ; \epsilon)$ in $\mathfrak{N}_\mathfrak{b}$ is 1-1 and order preserving. From this it follows that B is weakly complete relative to $\mathfrak{N}$ if and only if it is weakly complete relative to $\mathfrak{N}_\mathfrak{b}$.

Theorem 1 differs from Goldstine's theorem only in that $\mathfrak{N}$ replaces a class of partitions.

THEOREM 1. *B is reflexive if and only if it is weakly complete relative to $\mathfrak{N}$ and if and only if it is weakly complete.*

The known proof that reflexivity implies weak completeness is quite straightforward. If B is weakly complete relative to $\mathfrak{N}$, let $\mathfrak{b}_0$ be any point of B^{**}, let $N = N(\mathfrak{b}_0 ; \beta_1, \cdots, \beta_k ; \epsilon)$ be any weak* neighborhood of $\mathfrak{b}_0$, and consider the set of equations $\beta_i(b) = \mathfrak{b}_0(\beta_i)$ for $i = 1, \cdots, k$. For all choices of $t_1, \cdots, t_k$,

$$| \Sigma_{i \leqq k} t_i \mathfrak{b}_0(\beta_i) | = | \mathfrak{b}_0(\Sigma_{i \leqq k} t_i \beta_i) | \leqq \| \mathfrak{b}_0 \| \, \| \Sigma_{i \leqq k} t_i \beta_i \|.$$

Lemma 1 asserts for any $\delta > 0$ the existence of a point in B which satisfies these k equations and is of norm $< \| \mathfrak{b}_0 \| + \delta$. Hence there is a point b_N in B such that $\| b_N \| \leqq \| \mathfrak{b}_0 \|$ while $| \mathfrak{b}_0(\beta_i) - \beta_i(b_N) | < \epsilon$ for $i = 1, \cdots, k$. Define $\mathfrak{b}_N$ in B^{**} by the equations $\mathfrak{b}_N(\beta) = \beta(b_N)$ for every β in B^*; then $\mathfrak{b}_N \epsilon N$ so weak* $-\lim_N \mathfrak{b}_N = \mathfrak{b}_0$; i.e., $\mathfrak{b}_0(\beta) = \lim_N \mathfrak{b}_N(\beta) = \lim_N \beta(b_N)$ for every β in B^*. Since B is weakly complete relative to $\mathfrak{N}_\mathfrak{b}$, there is a b_0 in B such that $\mathfrak{b}_0(\beta) = \lim_N \beta(b_N) = \beta(b_0)$ for every β in B^*, so B is reflexive.

Note that in the course of this proof we have demonstrated, without use of any form of integration,

COROLLARY 1. *If B_0 is the image of B in B^{**} under the usual mapping and if E is any sphere about the origin in B^{**}, then E is the weak* closure of $E \cdot B_0$.*

If $\mathfrak{N}'$ is the set of neighborhoods $N(\theta; \beta_1, \cdots, \beta_k ; \epsilon)$ where the β_i are all chosen from some fundamental set[5] E_0 in B^*, it is clear that $\mathfrak{N}'$ and $\mathfrak{N}$ are equivalent neighborhood systems in any sphere about the origin in B^{**}; hence, if $\| b_x \| \leqq K$ for every x, $\lim_x b_x(\beta) = \mathfrak{b}_0(\beta)$ for every β in B^* if and only if $\lim_x b_x(\beta) = \mathfrak{b}_0(\beta)$ for every β in E_0. Repetition of the pertinent parts of the above proof gives

COROLLARY 2. *Weak completeness relative to $\mathfrak{N}'$ is necessary and sufficient for reflexivity of B.*

Following Tukey [7] let S^λ be a set of power $\aleph_\lambda$; the *stack* Σ^λ on the base S^λ is the directed set of all finite subsets σ of S^λ where $\sigma > \sigma'$ means $\sigma \supset \sigma'$. A

[4] θ will be used for the zero element of any linear space.

[5] A set $E \subset B$ is fundamental if B is the smallest closed linear set containing E.

subspace[6] B' of a Banach space B is called *λ-separable* if and only if a dense subset of power $\leqq \aleph_\lambda$ exists in B'.

COROLLARY 3. *If B^* is λ-separable, B is reflexive if and only if it is weakly complete relative to Σ^λ.*

All that need be proved here is that if B^* is λ-separable and $\mathfrak{b} \in B^{**}$, then for each σ in Σ^λ b_σ in B can be so chosen that $\| b_\sigma \| < \| \mathfrak{b} \| + 1$ while $\lim_\sigma \beta(b_\sigma) = \mathfrak{b}(\beta)$ for every β in B^*. Choose β_s, $s \in S^\lambda$, to be a fundamental set in B^*; by Lemma 1 there is for each σ a b_σ such that $\| b_\sigma \| < \| \mathfrak{b} \| + 1$ while b_σ satisfies the set of equations $\beta_s(b_\sigma) = \mathfrak{b}(\beta_s)$ for each s in σ. Then $\beta_s(b_\sigma) = \mathfrak{b}(\beta_s)$ if $s \in \sigma$, so $\lim_\sigma \beta_s(b_\sigma) = \mathfrak{b}(\beta_s)$ for every s; since the β_s form a fundamental set in B and $\| b_\sigma \| < \| \mathfrak{b} \| + 1$ for every σ, $\lim_\sigma \beta(b_\sigma) = b(\beta)$ for every β in B^*.

Note that in the case $\lambda = 0$ this asserts that weak completeness relative to Σ^0 implies reflexivity of B if B^* is separable in the ordinary sense; i.e., 0-separable. This is equivalent to Goldstine's corollary since weak completeness relative to Σ^0 and weak completeness relative to the directed set of positive integers (ordered by magnitude) are equivalent.

Let us turn now to some different characterizations of reflexivity. An example in l_1 shows that extension of the property described by Lemma 1 to infinite sets of β's and c's is not always possible. We shall show that the possibility of extension is closely connected with reflexivity. Let S be any set and B any Banach space, let β_s in B^* and c_s, real numbers, be given; for each subset S' of S let $E_{S'} = \{b \mid \beta_s(b) = c_s \text{ for each } s \text{ in } S'\}$; Lemma 1 gives an expression for $r(E_\sigma)$ for each finite subset σ of S.

THEOREM 2. *If B^* contains a total[7] set of power $\aleph_\lambda$, and if choosing β_s and c_s for each s in S^λ so that $\sup_\sigma r(E_\sigma) < \infty$ implies that E_S is not empty, then B is reflexive.*

Let $\{\beta_s \mid s \in S^\lambda\}$ be a total set in B^*, let $\mathfrak{b}$ be any element of B^{**}, and let $c_s = \mathfrak{b}(\beta_s)$. Then, as before, $r(E_\sigma) \leqq \| \mathfrak{b} \|$ for every σ, so there is a b in B such that $\beta_s(b) = \mathfrak{b}(\beta_s)$ for every s. If β is any other element of B^*, adding the equation $\beta(b) = \mathfrak{b}(\beta)$ to the given set gives a new system which has a solution b', but $\beta_s(b') = \beta_s(b)$ for every s and the set of all β_s is total, so $b' = b$. Therefore $\mathfrak{b}(\beta) = \beta(b)$ for every β in B^*.

This theorem is a consequence of the next result to be proved but the proof in this case is much simpler. It is known [6] that if $\lambda = 0$, the condition on β_s and c_s that $\sup_\sigma r(E_\sigma) < \infty$ implies that E_{S^0} is not empty is equivalent to weak sequential compactness of the unit sphere in B; this and other considerations suggest the conditions of Theorem 3. A set B' in B is *weakly λ-compact*[8]

[6] A subspace is a closed linear subset of B.

[7] A subset C of B^* is total if $\| b \| = 0$ whenever $\beta(b) = 0$ for every β in C.

[8] The referee remarks that Alexandroff and Urysohn in the Amsterdam Proceedings of 1929 use the phrase "Y has $[0, \lambda]$ as an interval of compactness" in the sense that "Y is λ-compact" has here; clearly this and conditions (12) and (13) are quantitative statements derived in the natural way from the usual conditions for "bicompactness" of a space in terms of open sets, closed sets, and convergence.

if every covering of B' by $\aleph_\lambda$ sets open in the weak topology in B' contains a finite subcovering of B'. If X is a directed set and B' is a subset of B, B' contains inner limits relative to X if $\| b_x \|$ bounded and $b_x \in B'$ for all x imply that a b_0 in B' exists for which $\liminf_x \beta(b_x) \leqq \beta(b_0) \leqq \limsup_x \beta(b_x)$ for every β in B^*.

THEOREM 3. *For each ordinal* $\lambda \geqq 0$ *the following conditions on a Banach space* B *are equivalent*:

(1) *Every* λ*-separable subspace of* B *is reflexive.*

(2) B *contains inner limits relative to* Σ^λ.

(3) *If for each* s *in* S^λ β_s *and* c_s *are so chosen that* $M = \sup_\sigma r(E_\sigma) < \infty$, *then there is a point in* E_{S^λ} *of norm* M.

(4) *Under the same conditions on* β_s *and* c_s, $\sup_\sigma r(E_\sigma) = r(E_{S^\lambda})$.

(5) *Under the same conditions on* β_s *and* c_s, E_{S^λ} *is not empty.*

(6) *If* $\mathfrak{b} \in B^{**}$ *and if for each* s *in* S^λ $\beta_s \in B^*$, *there is a* b *in* B *such that* $\beta_s(b) = \mathfrak{b}(\beta_s)$ *for every* s.

(7) *Every* λ*-separable subspace of* B^* *is weak* closed.*

(8) *Every* λ*-separable subspace of* B^* *is reflexive.*

(9) *The unit sphere in* B *is weakly* λ*-compact.*

(10) *If* C_s, $s \in S^\lambda$, *are closed convex subsets of the unit sphere in* B *such that* $\prod_{s \in \sigma} C_s$ *is not empty for any* σ, *then* $\prod_{s \in S^\lambda} C_s$ *is not empty.*

(11) *If* Ω_λ *is the first ordinal of power* $\aleph_\lambda$ *and if for each* $\mu \leqq \Omega_\lambda$ *the sets* C_ν, $\nu < \mu$, *are convex, closed, non-empty subsets of the unit sphere in* B *such that* $C_{\nu'} \subset C_\nu$ *if* $\nu' > \nu$, *then* $\prod_{\nu<\mu} C_\nu$ *is not empty.*

$(1 \to 2)$ If b_σ is defined for every σ in Σ^λ so that $\| b_\sigma \| \leqq K$ for every σ, let P be the smallest subspace of B containing all b_σ; then P is λ-separable. Since $\limsup_\sigma \pi(b_\sigma) \leqq K \| \pi \|$ for every π in P^*, by the Hahn-Banach theorem there is a $\mathfrak{p}$ in P^{**} such that $\mathfrak{p}(\pi) \leqq \limsup_\sigma \pi(b_\sigma)$ for every π in P^*. Since $\liminf_\sigma \pi(b_\sigma) = -\limsup_\sigma -\pi(b_\sigma) \leqq -\mathfrak{p}(-\pi) = \mathfrak{p}(\pi)$, $\liminf_\sigma \pi(b_\sigma) \leqq \mathfrak{p}(\pi) \leqq \limsup_\sigma \pi(b_\sigma)$ for every π in P^*. Since P is reflexive there is a p in P with $\mathfrak{p}(\pi) = \pi(p)$ for all π; since each β in B^* defines a unique π_β in P^*, $\liminf_\sigma \beta(b_\sigma) \leqq \beta(p) \leqq \limsup_\sigma \beta(b_\sigma)$.

$(2 \to 3)$ If β_s and c_s are given for each s in S^λ, let $| \sigma |$ be the number of elements of the finite subset σ of S^λ and by Lemma 1 take b_σ in E_σ so that $\| b_\sigma \| < r(E_\sigma) + 1/| \sigma |$. By (2) there is a b such that $\liminf_\sigma \beta(b_\sigma) \leqq \beta(b) \leqq \limsup_\sigma \beta(b_\sigma)$ for every β in B^*, so $\| b \| \leqq \limsup_\sigma \| b_\sigma \| \leqq M$. For each s, $\beta_s(b_\sigma) = c_s$ if $s \in \sigma$, so $\lim_\sigma \beta_s(b_\sigma) = c_s$ for every s; hence $\beta_s(b) = c_s$ for every s.

$(3 \to 4 \to 5)$ Obvious.

$(5 \to 6)$ By the argument used in Theorems 1 and 2, if we choose $c_s = \mathfrak{b}(\beta_s)$.

$(6 \to 7)$ If E is a λ-separable subspace of B^*, let $\{\beta_s \mid s \in S^\lambda\}$ be a dense (in norm) subset of E and let $\beta_0 \in B^* - E$. Then [1, p. 57] there is a $\mathfrak{b}_0$ in B^{**} such that $\mathfrak{b}_0(\beta) = 0$ if $\beta \in E$ while $\mathfrak{b}_0(\beta_0) = 1$. By (6) there is a b in B for which $\beta_s(b) = 0$ for every s while $\beta_0(b) = 1$; since the β_s are dense in E, $\beta(b) = 0$ for

every β in E, so the weak* neighborhood $N(\beta_0\,; b; 1)$ is disjoint from E and E is weak* closed.

$(7 \to 8)$[9] If E is a λ-separable subspace of B^*, by (7) every subspace of E is weak* closed in B^*. Let $E_0 = \{b \mid \beta(b) = 0$ for every β in $E\}$; then $E = \{\beta \mid \beta(b) = 0$ for every b in $E_0\}$ and, if A is the factor space[10] B/E_0, E is equivalent to A^* under the transformation associating β in E with α_β in A^* if $\alpha_\beta(a) = \beta(b)$ for some (or all) b in a. If A' is a subspace of A^* and if $E' = \{\beta \mid \alpha_\beta \in A'\}$, if $\alpha_0 \in A^* - A'$, and if β_0 is the image of α_0, then there is a b_0 in B such that $\beta(b_0) = 0$ if $\beta \in E'$ while $\beta_0(b_0) = 1$. b_0 is in some a_0 of A, so $\alpha(a_0) = 0$ if $\alpha \in A'$ while $\alpha_0(a_0) = 1$; hence A' is weak* closed in A^*. Since weak* and regular closure are the same for linear subsets of A^*, by [1], p. 132, A is reflexive so E is reflexive.

$(8 \to 1)$ B is equivalent to a subspace of B^{**}; by the implication $(1 \to 8)$ already established, reflexivity of every λ-separable subspace of B^* implies the same for B^{**} and hence for B.

To deal with (9) note that λ-compactness of a topological space Y is equivalent to either of two conditions.

(12) If for each s in S^λ F_s is a closed subset of Y and if $\prod_{s \in \sigma} F_s$ is not empty for any σ, then $\prod_{s \in S^\lambda} F_s$ is not empty.

(13) Every function on Σ^λ to Y has a cluster point.[11]

(12) is adapted from the standard formulation of compactness in terms of closed sets. If for each σ in Σ^λ a point y_σ is given, let F_σ be the closure of the set $\{y_{\sigma'} \mid \sigma' > \sigma\}$; then (12) implies the existence of a cluster point y_0 in all F_σ. If F_s, $s \in S$, are closed sets of Y such that $\prod_{s \in \sigma} F_s$ is never empty, choose y_σ in $\prod_{s \in \sigma} F_s$; then $\{y_{\sigma'} \mid \sigma' > \sigma\}$ is contained in F_s if $s \in \sigma$, so any cluster point y_0 of the y_σ is in every F_s.

$(13 \to 2)$ If b_σ, $\sigma \in \Sigma^\lambda$, are in the unit sphere of B, any weak cluster point b_0 of the b_σ will satisfy the equations $\liminf_\sigma \beta(b_\sigma) \leqq \beta(b_0) \leqq \limsup_\sigma \beta(b_\sigma)$ for every β in B^*.

$(1 \to 13)$ If for every σ in Σ^λ b_σ is a point of the unit sphere in B, let P be the smallest subspace of B which contains all b_σ. By (1) P is reflexive; hence its unit sphere is compact in the weak topology in P[see 6]. Stated in full this means that there is a p in P such that for every $\pi_1, \cdots, \pi_k$ in P^*, every $\epsilon > 0$, and every σ in Σ^λ there is a σ' in Σ^λ such that $\sigma' > \sigma$ and $| \pi_i(p) - \pi_i(b_{\sigma'}) | < \epsilon$ for $i = 1, \cdots, k$. Hence for every $\beta_1, \cdots, \beta_k$, $\epsilon > 0$, and σ in Σ^λ a σ' exists such that $\sigma' > \sigma$ and $| \beta_i(p) - \beta_i(b_{\sigma'}) | = | \pi_{\beta_i}(p) - \pi_{\beta_i}(b_\sigma) | < \epsilon$ for $i = 1, \cdots, k$. That is, p is a cluster point of b_σ in the weak topology in B.

Clearly $(9 \to 10 \to 3)$ and $(10 \to 11 \to 10)$.

[9] This step of the proof is due to Šmulian [6].

[10] The factor space B/E_0 is the space of cosets a of B, where b and b' both belong to the same a if $b - b' \in E_0$; $\| a \| = r(a)$.

[11] A point y_0 is a cluster point of a function f defined on a directed set X to a topological space Y if and only if for every neighborhood N of y_0 and every x in X there is an x' in X such that $x' > x$ and $f(x') \in N$. This is equivalent to the statement that y_0 is a cluster point of f if y_0 is in the closure of all the sets $\{f(x') \mid x' > x\}$ for every choice of x in X.

Let $R(\lambda)$ be the class of those Banach spaces with one (and hence all) of the properties (1)–(11).

COROLLARY 4. *If $B^{(n)}$ is defined by induction from $B^{(0)} = B$, $B^{(n+1)} = (B^{(n)})^*$, all $B^{(n)}$ are in $R(\lambda)$ if and only if some $B^{(n_0)}$ is in $R(\lambda)$.*

A Banach space B' is a *linear image* of a Banach space B if there is a linear operator U on B whose range is all of B'. If $E_0 = \{b \mid U(b) = \theta\}$, then B' is isomorphic to the space $A = B/E_0$ under the transformation associating a in A with b' in B' if $U(b) = b'$ for some (or all) b in a.

THEOREM 4. *B is in $R(\lambda)$ if and only if every linear image of B is in $R(\lambda)$.*

Since B is a linear image of B, the condition is necessary. Since the properties (1)–(11) are invariant under isomorphism, all that need be proved is that if E_0 is a subspace of B and if B has property (3), then $A = B/E_0$ has property (6). Take α_s in A^* and $\mathfrak{a}$ in A^{**}; define β_s in B^* by $\beta_s(b) = \alpha_s(a)$ if $b \in a$. Then if $c_s = \mathfrak{a}(\alpha_s)$, $| \Sigma_{s \epsilon \sigma} t_s c_s | \leqq \| \mathfrak{a} \| \| \Sigma_{s \epsilon \sigma} t_s \alpha_s \| \leqq \| \mathfrak{a} \| \| \Sigma_{s \epsilon \sigma} t_s \beta_s \|$ so, by (3), there is a b in B such that $\beta_s(b) = \mathfrak{a}(\alpha_s)$ for every s. Let a be the element of A to which b belongs; then $\alpha_s(a) = \beta_s(b) = \mathfrak{a}(\alpha_s)$ for every s.

COROLLARY 5. *Every λ-separable linear image of a space in $R(\lambda)$ is reflexive.*

Gantmakher and Šmulian note that sequential weak compactness of the unit sphere in B implies that each $\mathfrak{b}$ in B^{**} is weak* sequentially continuous as a function on B^*; i.e., that if $\beta_n(b) \to \beta_0(b)$ for every b in B, then $\mathfrak{b}(\beta_n) \to \mathfrak{b}(\beta_0)$ for every $\mathfrak{b}$ in B^{**}. Generalizing this to other cardinal numbers suggests the following property: $B \in P(\lambda)$ means that if β_σ is given for each σ in Σ^λ so that $\| \beta_\sigma \| \leqq K$ for all σ and $\lim_\sigma \beta_\sigma(b) = \beta_0(b)$ for every b in B, then $\lim_\sigma \mathfrak{b}(\beta_\sigma) = \mathfrak{b}(\beta_0)$ for every $\mathfrak{b}$ in B^{**}.

THEOREM 5. *Weak λ-compactness of the unit sphere in B implies that $B \in P(\lambda)$; if $B \in P(\lambda)$, B^* is weakly complete relative to Σ^λ.*

If B is weakly λ-compact and if $\mathfrak{b}$, β_0, and β_σ have the given properties, by (6) of Theorem 3 there is a b in B such that $\mathfrak{b}(\beta_0) = \beta_0(b)$ and $\mathfrak{b}(\beta_\sigma) = \beta_\sigma(b)$ for all σ; hence $\lim_\sigma \mathfrak{b}(\beta_\sigma) = \mathfrak{b}(\beta_0)$. If $B \in P(\lambda)$ and β_σ is given for σ in Σ^λ so that $\| \beta_\sigma \| \leqq K$ and $\lim_\sigma \mathfrak{b}(\beta_\sigma)$ exists for every $\mathfrak{b}$ in B^{**}, by (13) there is a β_0 in B^* which is a cluster point of β_σ in the weak* topology. Hence $\lim_\sigma \beta_\sigma(b) = \beta_0(b)$ for each b in B; since $B \in P(\lambda)$, $\lim_\sigma \mathfrak{b}(\beta_\sigma) = \mathfrak{b}(\beta_0)$ and B^* is weakly complete.

Note that the proof of Corollary 3 can be used to show that if B^* is λ-separable and in $P(\lambda)$, B^* (and hence B) is reflexive (the argument proves that $B_0 = B^{**}$). Theorems 1 and 5 show that if $B \in P(\lambda)$ and B^{**} is λ-separable, then B is reflexive. The next lemma allows us to prove a result including both of these. Call H a hyperplane in B if $H = \{b \mid \beta(b) = c\}$ where $\| \beta \| \neq 0$ and c is any real number.

LEMMA 2. *If B is λ-separable and H is any hyperplane in B^*, any point in the weak* closure of H is the weak* limit of a bounded function on Σ^λ to H.*

If H is the hyperplane, it suffices to consider the case in which the point under consideration is not in H; since every hyperplane of the form $\{\beta \mid \beta(b) = c\}$ for some b and c is weak* closed, it suffices to consider a hyperplane H of the form $\{\beta \mid \mathfrak{b}(\beta) = c\}$ where $\mathfrak{b} \epsilon' B_0$, the image of B under the usual mapping into B^{**}; by the uniformity of the weak* topology in B* it suffices to take $c \neq 0$ and θ for the point. If $N = N(b_1, \cdots, b_k; \epsilon)$ is any weak* neighborhood of θ, Lemma 1 asserts the existence of a solution of the set of equations $\mathfrak{b}(\beta) = c$, $\beta(b_i) = 0$ for $i = 1, \cdots, k$, whose norm is arbitrarily near sup $| t_0 c |/|| t_0 \mathfrak{b} - \Sigma_{i=1}^k t_i \mathfrak{b}_i ||$ (where $\mathfrak{b}_i$ is the image of b_i in B^{**}). This $\leqq c/K$, where K is the distance from $\mathfrak{b}$ to B_0 so θ is in the weak* closure of a bounded subset of H. By arguments used before when B is λ-separable, it follows that β_σ in H can be so chosen for σ in Σ^λ that $|| \beta_\sigma || < 1 + c/K$ and that $w^*\text{-lim}_\sigma \beta_\sigma = \theta$.

Note that this shows incidentally that every point in B^* is the weak* limit of such a function, provided only that $\mathfrak{b} \epsilon' B_0$.

Theorem 6. *If B is λ-separable and in $P(\lambda)$, then B is reflexive.*

Take a $\mathfrak{b}$ not in B_0 and any c and consider the hyperplane $H = \{\beta \mid \mathfrak{b}(\beta) = c\}$; by the preceding lemma and the remark, if $\beta_0 \epsilon B^*$ there is a bounded function on Σ^λ with values in H converging in the weak* topology to β_0; hence, by the property $P(\lambda)$, the same function converges weakly to β_0 so β_0 is in the weak closure of H. But H is weakly closed so $H = B^*$. This can not happen if $|| \mathfrak{b} || \neq 0$; hence $\mathfrak{b}$ can not be chosen in $B^{**} - B_0$, or B is reflexive.

This leaves unanswered the more difficult question raised by Gantmakher and Šmulian for $\lambda = 0$: Is it true that $B \epsilon P(\lambda)$ if and only if $B \epsilon R(\lambda)$? Proving this is equivalent to showing that the property $P(\lambda)$ holds for every subspace of B if it holds for B.

Corollary 6. *$B \epsilon R(\lambda)$ if and only if every subspace of B is in $P(\lambda)$ or if every λ-separable subspace of B is in $P(\lambda)$.*

The condition that $B \epsilon P(\lambda)$ can be replaced by a stronger condition: Every weak* cluster point of a bounded function on Σ^λ to B^* is a weak cluster point; B can easily be shown to have this property if and only if $B \epsilon R(\lambda)$ because this property can easily be proved to hold in any subspace of a space for which it holds. A similar (and equivalent) condition can be formulated in terms of inner limits and their weak* analogues.

Theorem 7. *$B \epsilon P(\lambda)$ if and only if every linear image of B is in $P(\lambda)$.*

Using the notation of Theorem 4, take α_σ in A^* converging to α_0 in the weak* topology and $\mathfrak{a}$ in A^{**}; for each α in A^* define $T\alpha$ in B^* by $T\alpha(b) = \alpha(a)$ if $b \epsilon a$. Let $\beta_s = T\alpha_s$, $\beta_0 = T\alpha_0$ and let $\mathfrak{b}$ be any element of B^{**} such that $\mathfrak{b}(T\alpha) = \mathfrak{a}(\alpha)$ for all α. Then each b is in some a so $\lim_\sigma \beta_\sigma(b) = \lim_\sigma \alpha_\sigma(a) = \alpha_0(a) = \beta_0(b)$, so, since $B \epsilon P(\lambda)$, $\lim_\sigma \mathfrak{b}(\beta_\sigma) = \mathfrak{b}(\beta_0)$. Hence $\lim_\sigma \mathfrak{a}(\alpha_\sigma) = \mathfrak{a}(\alpha_0)$, and $A = B/E_0 \epsilon P(\lambda)$. Since the property $P(\lambda)$ is obviously invariant under isomorphism, the theorem is proved.

Notice that both this theorem and Theorem 4 could be proved in an essentially equivalent way by the use of adjoint operations instead of factor spaces.

Still unsolved is the question whether $R(0)$ is the class of reflexive spaces. Clearly $R(\lambda) \supset R(\mu)$ if $\lambda \leqq \mu$; no other relations are known.

Bibliography

1. S. Banach, *Théorie des opérations linéaires*, Monografje Matematyczne, I, Warsaw, 1932.
2. G. Birkhoff, *Moore-Smith convergence in general topology*, Annals of Mathematics, vol. 38(1937), pp. 39–56.
3. H. H. Goldstine, *Weakly complete Banach spaces*, Duke Mathematical Journal, vol. 4(1938), pp. 125–131.
4. S. Kakutani, *Weak topology and regularity of Banach spaces*, Imperial Academy of Tokyo, Proceedings, vol. 15(1939), pp. 169–173.
5. E. H. Moore and H. L. Smith, *A general theory of limits*, American Journal of Mathematics, vol. 44(1922), pp. 102–121.
6. V. Šmulian, *Several notes in the Comptes Rendus (Doklady) de l'Academie des Sciences de l'URSS*, in particular, vol. 18(1938), pp. 405–407, and vol. 17(1937), pp. 91–94 (with L. Gantmakher).
7. J. W. Tukey, *Convergence and uniformity in topology*, Annals of Mathematics Studies, No. 2, Princeton, 1940.

University of Illinois.